



(19)



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 288 881**

(51) Int. Cl.:

C12N 15/82 (2006.01)

C12N 15/29 (2006.01)

C07K 14/415 (2006.01)

A01H 5/00 (2006.01)

(12)

TRADUCCIÓN DE PATENTE EUROPEA

T3

(86) Número de solicitud europea: **00977111 .4**

(86) Fecha de presentación : **10.11.2000**

(87) Número de publicación de la solicitud: **1228228**

(87) Fecha de publicación de la solicitud: **07.08.2002**

(54) Título: **Genes de enanismo del sorgo y métodos de uso.**

(30) Prioridad: **12.11.1999 US 165176 P**

(45) Fecha de publicación de la mención BOPI:
01.02.2008

(45) Fecha de la publicación del folleto de la patente:
01.02.2008

(73) Titular/es:
PIONEER HI-BRED INTERNATIONAL, Inc.
800 Capital Square, 400 Locust Street
Des Moines, Iowa 50309, US
The Curators of the University of Missouri

(72) Inventor/es: **Johal, Gurmukh S. y**
Multani, Dilbag S.

(74) Agente: **Carvajal y Urquijo, Isabel**

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

DESCRIPCIÓN

Genes de enanismo del sorgo y métodos de uso.

5 **Campo de la invención**

La presente invención se relaciona con la manipulación genética de organismos, particularmente plantas, con genes de control de crecimiento y desarrollo. La invención además se relaciona con genes que controlan el crecimiento, incluyendo las formas homólogas y mutantes, las proteínas codificadas de estos y las plantas transformadas con estos genes.

Antecedentes de la invención

Las plantas enanas han tenido un gran impacto en la agricultura. Las variedades enanas del trigo se utilizan ampliamente en Norte América debido al potencial reducido para el encamado de las cosechas y rendimientos altos. Los árboles frutales enanos también se utilizan extensamente y permiten que los agricultores produzcan más fruta por acre, de tal manera que aumenta el potencial económico de la producción. Hay otros beneficios que se pueden obtener con el uso de plantas de cultivos enanas y árboles frutales enanos incluyendo, reducciones en las cantidades requeridas de pesticidas y fertilizantes, densidades altas de plantación y reducidos costos de trabajo.

Debido a las actuales tendencias del incremento de la población humana y la disminución del área de tierra apropiada para la agricultura, el incremento de la productividad agrícola es, y continuará siendo, un desafío de suprema importancia. Las plantas y árboles frutales enanos han sido y continuarán siendo importantes componentes de nuestro sistema de producción agrícola. El uso creciente de las plantas de cultivos enanas y árboles frutales enanos puede ayudar a resolver las demandas de producción agrícola del futuro. Sin embargo, las variedades enanas comercialmente aceptables no están disponibles para todos los cultivos.

Además del uso de las plantas enanas para controlar la altura de una planta, los productos químicos sintéticos se aplican rutinariamente a ciertas especies de plantas económicamente importantes para reducir el crecimiento. Los reguladores de crecimiento de plantas conocidos como retardantes de crecimiento se utilizan para reducir la elongación del tallo en una variedad de cultivos incluyendo algodón, vinos de uvas, árboles frutales, maní, trigo y ornamentales tales como azaleas, crisantemos, hortensias, poinsetias y muchas plantas de ornato. Todos los retardantes de crecimiento utilizados comúnmente son inhibidores de la biosíntesis de la giberilina y limitan el crecimiento del tallo o brote, reduciendo la elongación. En los Estados Unidos, el retardante de crecimiento más ampliamente usado es el cloruro de mepiquat, el cual se registró para usar en algodón. Las ventajas atribuidas al uso del cloruro de mepiquat en algodón incluyen el incremento de la producción, desfoliación mejorada, tolerancia al estrés mejorada, madurez del cultivo más uniforme y la capacidad para cosechar más temprano. Previamente, el retardante de crecimiento, daminozida, se registró para usar en los Estados Unidos en manzanas, uvas y maní bajo las marcas comerciales ALAR y KYLAR pero se retiraron de uso en cultivos de alimentos debido a la preocupación por la salud humana. A pesar de las demandas de los productores agrícolas de un producto para reemplazar la daminozida, no existen retardantes de crecimiento registrados para utilizar en uvas, árboles frutales y maní en los Estados Unidos. La daminozida, sin embargo, sigue siendo ampliamente utilizada en ciertas especies de plantas, no-comestibles.

Descubrir los mecanismos moleculares que controlan los procesos de crecimiento de la planta tales como división celular y elongación celular probablemente ayudará, en el desarrollo de nuevas variedades de plantas con estatura reducida y nuevos métodos para reducir el crecimiento de la planta. Tales nuevas variedades de plantas y métodos pueden proporcionar a los agricultores y horticultores de alternativas del uso de productos químicos sintéticos, ambientalmente benignos, que retarden el crecimiento.

La elongación de las células vegetales y órganos es uno de los parámetros más críticos del crecimiento y desarrollo de la planta. La regulación de esta característica en las plantas, sin embargo, es un proceso bastante complicado, pues los factores externos e internos lo influyen. El estímulo externo más importante es ligero, con su efecto normalmente reprimible o negativo sobre la elongación celular (Quail, P.H. (1995) *Science* 268:675-680; Kende *et al.* (1997) *Plant Cell* 9:1197-1210). El control interno de elongación celular se media por un número de productos químicos, normalmente refiriéndose a los reguladores de crecimiento de la planta u hormonas (Kende *et al.* (1997) *Plant Cell* 9:1197-1210). Entre las hormonas vegetales clásicas, auxinas y giberelinas (GAs) ambas promueven la elongación celular mientras que se ha mostrado, que las citoquininas y ácido abscísico tienen un efecto negativo en la elongación celular (Kende *et al.* (1997) *Plant Cell* 9:1197-1210). Recientemente, otra clase de reguladores de crecimiento de la planta, llamado brasinoesteroides, se ha identificado que también promueve dramáticamente el crecimiento de la planta (Yokota, T. (1997) *Trends Plant Sci.* 2:137-143; Azpiroz *et al.* (1998) *Plant Cell* 10:219-230; Choe *et al.* (1998) *Plant Cell* 10:231-243). Sin embargo, los mecanismos por los cuales las hormonas de las plantas actúan, tanto individualmente como en combinación, para controlar la elongación celular, permanecen confusos.

Una manera de ganar una comprensión de los mecanismos que median la elongación celular es estudiar los mutantes en los cuales este aspecto del crecimiento de la planta se comprometen (Klee *et al.* (1991) *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 42:529-551). Numerosos mutantes han sido identificados a través de más especies de plantas, incluyendo maíz, en la cual se han caracterizado, más de 25 mutaciones genéticas monofactoriales que afectan la estatura de la planta (Coe *et al.* (1988) In: *Corn & Corn Improvement*. G. F. Sprague (Ed.) Madison, WI; She-

ridan, W.F. (1988) Annu. Rev. Genet. 22:353-385). Estos mutantes enanos se consideran que son GA relacionados, principalmente porque GA es la única fitohormona, cuyo papel en la regulación de la altura en maíz se ha establecido convincentemente (Phinney *et al.* (1985) Curr. Top. Plant Biochem. Physiol. 4:67-74; Fujioka *et al.* (1988) Proc. Natl. Acad. Sci. USA 85:9031-9035). Ambos tipos de mutantes, GA sensible y GA no-sensible, se han encontrado

- 5 en esta colección de mutantes del maíz. Mientras que para un número de mutantes GA-sensibles se han clonado y encontrado que se involucran en la biosíntesis de GA (Bensen *et al.* (1995) Plant Cell 7:75-84; Winkler *et al.* (1995) Plant Cell 7: 1307-1317), nada se conoce acerca de la naturaleza de los defectos en los mutantes GA no-sensibles del maíz.
- 10 Un tipo de mutantes enanos GA no-sensibles que ha recibido mucha atención de los genetistas y criadores del maíz se llama braquítico. Estos enanos se caracterizan por entrenudos de longitud reducida considerablemente, relativo al tipo-salvaje, sin que tenga ningún efecto sobre le tamaño o número de otros órganos, incluyendo las hojas, mazorca y borla (Kempton, J.H. (1920) J. Hered 11:111-115). Hay tres conocidas mutaciones braquíticas en maíz, *br1*, *br2* y *br3*, que son resécivas (Coe *et al.* (1988) In: Corn & Corn Improvement, G. F. Sprague (Ed.) Madison, WI; Sheridan, W.F.
- 15 (1988) Annu. Rev. Genet. 22:353-385). Debido al interés comercial en *br2* para mejorar la productividad de la planta (Pendleton *et al.* (1961) Crop Sci. 1:433-435; Duvick, D.N. (1977) Maydica 22:187-196; Djisbar *et al.* (1987) Maydica 32:107-123; Russel, W.A. (1991) Adv. Agron. 46:245-298), esta enana ha sido la más caracterizada. Dependiendo de los antecedentes genéticos, las plantas resécivas homocigóticas para *br2* son 30-70% más pequeñas que sus hermanas normales. Esta reducción en altura de una planta es exclusivamente debido a una reducción de la longitud de los
- 20 entrenudos del tallo (tronco). Además de ser enanos, los mutantes *br2* crecen bajo condiciones de invernadero con frecuencia sufren de “buggy whip”, una condición como una enfermedad en la cual las hojas desplegadas en el espiral experimentan necrosis y permanecen pegadas la una con la otra. Esta condición con frecuencia resulta en la muerte del ápice de crecimiento de la planta.

- 25 Aunque la característica de enanismo en maíz se ha estudiado extensamente genética y molecularmente, tiene todavía que ser explotada con éxito en el cultivo de logros en esta planta de cultivo. En contraste, los mutantes enanos del sorgo han contribuido significativamente al desarrollo de cultivos de los días modernos. El sorgo y maíz son ambos miembros de la familia gramínea (Poaceae o Gramineae) y de esta manera comparte muchas características incluyendo organización genómica y forma del cuerpo de la planta. Fuera de las cuatro mutaciones de enanismo
- 30 explotadas en sorgo, *dw3*, cuyo fenotipo de enanismo parece muy similar a aquel del *br2* en maíz, parece ser el más notable. Sin embargo, el único alelo *dw3* (*dw3-ref*) disponible hasta el momento tiene un serio problema que limita su valor agronómico. El fenotipo enano asociado con el alelo *dw3* es inestable, con una frecuencia de reversión a tipo-salvaje (alto) tan alto como cerca del 1% en ciertos antecedentes genéticos. La inestabilidad de este fenotipo enano, el mecanismo del cual ha eluido a genetistas del sorgo hasta el momento, no solo continúa desconcertando
- 35 los criadores de sorgo, sino también conduce algunas veces al rechazo, por otra parte de un promotor congénito o híbrido.

- Continuar con la demanda para la producción agrícola creciente, nuevos objetivos son necesarios para la ingeniería genética de plantas agrícolas para la mejora de las características agronómicas. Elucidar los mecanismos moleculares de la división y elongación celular proporcionará los nuevos objetivos para los científicos agrónomos para manipular.
- 40

Resumen de la invención

- Se proporcionan, las composiciones y los métodos para los genes de expresión que codifican la P-glicoproteínas de la invención en las plantas. En particular, la invención suministra las secuencias de nucleótidos del gen *Dw3* del sorgo. Las secuencias de la invención son útiles en convertir las plantas para la expresión de tejido-preferido o constitutivo y encontrar el uso en métodos para controlar el crecimiento de organismos, particularmente el crecimiento del tallo en las plantas. Por consiguiente la invención suministra:
- 45

- 50 - una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos seleccionada del grupo que consiste de:

- (a) una secuencia de nucleótidos publicada en la SEQ ID NO: 1;
- 55 (b) una secuencia de nucleótidos publicada en la SEQ ID NO: 2;
- (c) una secuencia de nucleótidos publicada en la SEQ ID NO: 3;
- (d) una secuencia de nucleótidos publicada en la SEQ ID NO: 7;
- 60 (e) una secuencia de nucleótidos publicada en la SEQ ID NO: 8;
- (f) una secuencia de nucleótidos que consiste de al menos 150 nucleótidos contiguos de la secuencia de nucleótidos publicada en cualquiera de (a)-(e), en donde dicha secuencia de nucleótidos codifica una P-glicoproteína que controla el crecimiento de la planta;
- 65 (g) una secuencia de nucleótidos que codifica la secuencia de aminoácidos publicada en la SEQ ID NO: 9;

ES 2 288 881 T3

(h) una secuencia de nucleótidos que codifica al menos 70 aminoácidos contiguos de la secuencia de aminoácidos publicada en la SEQ ID NO: 9, en donde dicha secuencia de nucleótidos codifica una P-glicoproteína que controla el crecimiento de la planta;

(i) una secuencia de nucleótidos que comprende al menos 80% de identidad a la secuencia publicada en la SEQ ID NO: 7, en donde dicha secuencia de nucleótidos codifica una P-glicoproteína que controla el crecimiento de la planta;

(j) una secuencia de nucleótidos que comprende al menos 80% de identidad a la secuencia publicada en la SEQ ID NO: 8, en donde dicha secuencia de nucleótidos codifica una P-glicoproteína que controla el crecimiento de la planta;

(k) una secuencia de nucleótidos que es complementaria a la secuencia de nucleótidos de cualquiera de (a)-(k); y

(l) una secuencia de nucleótidos que hibridiza bajo condiciones rigurosas a la secuencia de nucleótidos cualquiera de (a)-(e), o a una secuencia complementaria de estas, en donde dichas condiciones rigurosas comprenden la hibridación en una solución que comprende 50% de formamida, NaCl 1 M, 1% de SDS a 37°C y un lavado en 0.1 X SSC a 60°C, y dicha secuencia de nucleótidos codifica una P-glicoproteína que controla el crecimiento de la planta;

- un método para modificar el crecimiento de una planta, dicho método que comprende convertir una planta con una secuencia de nucleótidos que codifica una P-glicoproteína en donde dicha P-glicoproteína funciona para controlar el crecimiento de un organismo, dicha secuencia de nucleótidos ligada operablemente para la producción de transcripciones antisentido a un promotor capaz de conducir la expresión de dicha secuencia en dicha planta; en donde dicha secuencia de nucleótidos es como se define por la invención; y

- un método para la identificación de una planta de sorgo que tiene un fenotipo enano estable que comprende:

(a) obtención de una planta de sorgo enano que es un descendiente de una planta de sorgo enano *dw3* inestable.

(b) análisis del ADN genómico de dicha planta de sorgo enano para determinar si al menos una copia de un gen *dw3* carece de una inserción, o al menos de una porción de dicha inserción;

(c) selección de dicha planta de sorgo enano si dicha planta de sorgo enano contiene en su genoma al menos una copia de dicho gen *dw3* carente de dicha inserción, o carente de dicha porción de estas.

Se proporcionan los métodos para identificar las plantas que poseen un alelo mutante que es capaz de conferir un fenotipo mutante estable sobre un organismo. Tales métodos encuentran uso en la agricultura, particularmente en el cultivo de las plantas de cultivos enanos, particularmente las plantas de sorgo enanos.

Se proporcionan además los casetes de expresión que comprenden las secuencias de la invención. Adicionalmente se proporcionan las plantas transformadas, tejidos de planta, células de plantas y semillas de estas. Se proporcionan las proteínas aisladas codificadas por las secuencias de nucleótidos de la invención. Por consiguiente la invención también suministra:

- una proteína aislada que comprende un miembro seleccionado del grupo que consiste de:

(a) un polipéptido que comprende la secuencia de aminoácidos publicada en la SEQ ID NO: 9;

(b) un polipéptido que comprende al menos 70 aminoácidos contiguos de la secuencia de aminoácidos publicada en la SEQ ID NO: 9, en donde dicho polipéptido es una P-glicoproteína que controla el crecimiento de la planta;

(c) un polipéptido codificado por la secuencia de nucleótidos publicada en la SEQ ID NO: 8; y

(d) un polipéptido codificado por una secuencia de aminoácidos que comprende al menos 80% de identidad a la secuencia amino de (a), en donde dicho polipéptido es una P-glicoproteína que controla el crecimiento de la planta.

Descripción detallada de la invención

La presente invención se relaciona, en particular, para controlar el crecimiento del tallo en las plantas. Así, se proporcionan las plantas transformadas, células de las plantas, tejidos de planta y semillas que comprenden los ácidos nucleicos de la invención.

Las composiciones de la invención incluyen las secuencias de nucleótidos nativos para el gen de la P-glicoproteína del sorgo *Dw3* locus y la respectiva secuencia de aminoácidos para las P-glicoproteínas codificadas por consiguiente,

según se reivindica. Las secuencias de nucleótidos *Dw3* se publican en la SEQ ID NOS: 1-3 y 7-8. Las secuencias de nucleótidos o las secuencias antisentido correspondientes encuentran uso en la modulación de la expresión de una P-glicoproteína en una planta o célula de la planta. Es decir, las secuencias codificadas se pueden utilizar para incrementar la expresión mientras las secuencias antisentido se pueden utilizar para disminuir la expresión.

Las secuencias de la invención encuentran uso en los métodos de modificación del crecimiento de la planta. Hacia este fin, las secuencias de la invención se pueden utilizar en casetes de expresión o construcciones del nucleótido ligados operablemente a cualquiera de una variedad de promotores de la planta. Los aspectos de crecimiento de la planta que se puede impactar por los métodos de la invención incluyen, pero no se limitan a, altura de una planta; el tamaño, forma y número de células y órganos; velocidad de división celular; velocidad de elongación celular; la velocidad de crecimiento de la planta, sus órganos, tejidos y células; medida y localización de la iniciación del órgano; duración; y similares.

La invención revela los métodos para reducir el crecimiento de la planta que encuentra uso, como alternativas a la aplicación sintética, de productos químicos que retardan el crecimiento a las plantas. Estos métodos proporcionan alternativas ambientalmente seguras a los medios tradicionales de retardar la elongación del tallo o el crecimiento con productos químicos sintéticos. Algunas modalidades de la invención utilizan las plantas transformadas con promotores de tejido-preferido, particularmente promotores de tallo-preferido, ligados operablemente a las secuencias de nucleótidos de la invención.

Se proporcionan los métodos para reducir el crecimiento de una planta. Tales métodos involucran convertir las plantas con al menos una secuencia de nucleótidos de la invención. Las secuencias de nucleótidos se pueden utilizar en tanto la orientación sentido como antisentido para suprimir el nivel de una P-glicoproteína endógena que controla el crecimiento de una planta. Reduciendo el nivel en una planta de una P-glicoproteína, particularmente una que controla el crecimiento del tronco o tallo, una planta de estatura reducida, se puede lograr, una planta enana. Se pueden obtener por estos métodos, plantas enanas que tienen características agronómicas mejoradas. Tales características agronómicas mejoradas incluyen, pero no se limitan a, potencial reducido para el encamado de la cosecha, incremento de la eficiencia del uso del agua, ciclo biológico reducido, incremento de la eficiencia de la cosecha y incremento producción por unidad de área. Los métodos de la invención pueden eliminar la necesidad de injertar brotes de árboles frutales en rizomas de enanismo para producir árboles frutales enanos.

Los métodos de la invención encuentran uso en la producción de variedades enanas de plantas de cosecha. En una modalidad de la invención, una planta enana de arroz Basmati se produce transformando la planta con una secuencia de nucleótidos de la invención. El arroz Basmati, conocido por su fragancia aromática, espigado, granos alargados, y tiempo de cocción relativamente breve, es el tipo de arroz favorito de la mayoría de la gente en el sub-continente Índio. Mientras los cultivos enanos comercialmente aceptables han sido desarrollados para otros tipos de arroz, previos intentos para producir variedades de arroz Basmati comercialmente aceptables por métodos tradicionales del cultivo de la planta han fallado. Aunque se obtuvieron plantas enanas en dichos intentos, algunas de las características distintivas del grano que los consumidores esperaban en el arroz Basmati no se retuvieron en las plantas enanas. Los métodos de la invención proveen un medio de construcción de plantas enanas de arroz Basmati que produce un grano que posee las características deseadas por los consumidores.

Las deseadas plantas enanas de arroz Basmati se producen transformando una planta de arroz Basmati no-enana con una secuencia de nucleótidos de la invención ligada operablemente a un promotor que conduce la expresión en una planta. Mientras la elección del promotor depende del resultado deseado, los promotores preferidos son promotores de tejido preferido, particularmente promotores de tallo preferido. A través de cosupresión (supresión sentido) o supresión antisentido, tales plantas producen niveles reducidos de al menos una P-glicoproteína que controla el crecimiento de la planta de arroz Basmati, particularmente el crecimiento del tallo.

También se proporcionan los métodos para la identificación de una planta de sorgo con un fenotipo enano estable. Tales métodos encuentran uso en la agricultura, particularmente en el desarrollo de plantas de cosecha mejoradas. Los métodos se relacionan con un fenotipo mutante, de inserción-inducida. Por "fenotipo mutante, de inserción-inducida" se entiende un fenotipo mutante que se debe a la inserción de un nucleótido, o una secuencia de nucleótidos, en la secuencia de un gen de interés. Mientras que la invención no depende de un mecanismo genético particular para tal fenotipo mutante de inserción-inducida, la presencia de tal inserción dentro de un gen interrumpe típicamente la función tipo-salvaje normal del gen, o del producto del gen de estos. Mientras que los métodos de la invención no se unen por ningún tipo de inserción particular, tal inserción se puede deber a, por ejemplo, la inserción de un transposón o de un elemento transponible, o a la duplicación de una secuencia de nucleótidos tal como aquellos conocidos que ocurren, como resultado de la recombinación genética.

Preferiblemente, tal fenotipo de inserción-inducida es inestable de una generación a la siguiente. Es decir, autofecundación de una o más plantas similares que tienen el fenotipo de inserción-inducida resulta en al menos una individual dentro de la población de la progenie resultante que ha revertido al fenotipo de tipo-salvaje. Más preferiblemente, tal inestabilidad fenotípica, de una generación a la siguiente, se debe a la pérdida de al menos una porción de la inserción del gen de interés y que tal pérdida resulta en al menos una planta progenie, que ha revertido a un fenotipo de tipo salvaje. Los métodos de la presente invención involucran la identificación de un individuo con un fenotipo mutante estable dentro de tal población progenie.

Para identificar una planta que posee un alelo de un gen que confiere un fenotipo mutante estable, El ADN genómico de una planta mutante se analiza para determinar si al menos una copia del gen de interés carece de la inserción, o al menos una porción de estas. Generalmente, la planta mutante se selecciona de una población de progenie derivada de la autofecundación de una o más plantas que tienen fenotipo mutante, de inserción-inducida. Típicamente, en una población de tal progenie, revertientes tipo-salvaje también serán observadas, indicando que al menos una porción de la inserción se ha tomado por excisión del gen de interés. El ADN genómico de la planta mutante seleccionada se puede aislar y analizar para la ausencia de toda o una porción de la inserción por técnicas conocidas por aquellos de ordinaria habilidad en el oficio tal como, por ejemplo, Transferencia de Southern, análisis de Polimorfismo de longitud de fragmentos de restricción (RFLP) y amplificación de ADN por reacción en cadena de la polimerasa (PCR). Una vez una planta mutante carente de una porción de la inserción se identifica, la progenie de tal planta mutante se puede monitorear para verificar la estabilidad fenotípica. Si se desea, las generaciones subsiguientes también se pueden monitorear.

Un método de la invención involucra la identificación de una planta de sorgo con un fenotipo enano estable. Tal planta de sorgo posee en su genoma un alelo mutante estable del gen *Dw3*. Tal alelo mutante estable es capaz de conferir un fenotipo enano estable sobre una planta de sorgo y la secuencia de nucleótidos de un fragmento de tal alelo se publica en la SEQ ID NO: 2. Un método de la invención emplea el análisis RFLP utilizando Transferencia de Southern con una sonda derivada de las secuencias de nucleótidos de maíz *Br2*. Este método adicionalmente involucra amplificación PCR y el análisis de la secuencia de ADN para determinar la secuencia de nucleótidos del alelo mutante estable.

Las P-glicoproteínas de la invención abarcan todos los polipéptidos y las secuencias de nucleótidos de la invención tanto si o no tales polipéptidos poseen carbohidratos o cadenas que contienen carbohidratos unidos covalentemente.

Por “fenotipo mutante” se entiende cualquier fenotipo de tipo no salvaje, no-típico o no-estándar que ocurre como resultado de una alteración genética en el genoma de un organismo. Cuando se usa en referencia a las plantas y animales domesticados, un “fenotipo mutante” es cualquier fenotipo que es considerablemente diferente del fenotipo típico de la particular especie domesticada o variedad cultivada de la cual el fenotipo mutante se presenta.

Por “planta mutante” se entiende una planta que tiene un fenotipo mutante.

Por “alelo mutante” se entiende un alelo de un gen que es capaz de causar un “fenotipo mutante”.

Por “enana” se entiende atípicamente pequeña. Por “planta enana” se entiende una planta atípicamente pequeña. Generalmente, tal “planta enana” tiene una estatura o altura que se reduce de aquella de una planta típica por aproximadamente 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60% o más. Generalmente, pero no exclusivamente, tal planta enana se caracteriza por una longitud reducida del tronco, tallo o trunk cuando se compara con la planta típica.

Por “molécula nucleótido” se entiende una molécula compuesta de nucleótidos covalentemente unidos unos a otros. Los nucleótidos incluyen ambos ribonucleótidos y desoxiribonucleótidos. “Molécula nucleótido” abarca formas monocatenaria y bicatenaria de ambos ADN y ARN. Las “Moléculas nucleótido” pueden ser naturales, sintéticas o una combinación de ambos. La distribución lineal de los nucleótidos en una “molécula nucleótido” se refiere como a una “secuencia de nucleótidos” y a menos que se especifique de otra manera se presenta aquí de izquierda a derecha correspondiente a la dirección 5'-a-3'. Debido a la naturaleza complementaria de las hebras opuestas de una molécula nucleótido de doble cadena, una secuencia de nucleótidos de la invención adicionalmente abarca su secuencia antisentido complementaria.

Las composiciones de la invención incluyen secuencias de nucleótidos nativas para genes que codifican P-glicoproteínas de resistencia-multifármaco similar-gen-codificado, homólogas de P-glicoproteínas resistencia-multifármaco-similar-gen-codificado, secuencias antisentido, así como también fragmentos y variantes y fragmentos de estas. En particular, la presente invención proporciona las moléculas de ácido nucleico aisladas que comprenden las secuencias de nucleótidos que codifican las secuencias de aminoácidos mostradas en la SEQ ID NOS: 4 y 9.

Adicionalmente se proporcionan los polipéptidos que tienen una secuencia de aminoácidos codificados por una molécula de ácido nucleico descritos aquí, por ejemplo aquellos publicados en la SEQ ID NOS: 3 y 8, respectivamente según se reivindica.

La invención abarca composiciones de proteínas o ácido nucleicos aislados o considerablemente purificados. Una proteína o molécula de ácido nucleico “aislada” o “purificada”, o una porción activa biológicamente de estas, está considerablemente libre de otro material celular, o medio de cultivo cuando se produce por técnicas recombinantes, o considerablemente libres de precursores químicos u otros productos químicos cuando se sintetiza químicamente. Preferiblemente, un ácido nucleico “aislado” es libre de secuencias (preferiblemente una proteína que codifica las secuencias) que naturalmente guarda el flanco del ácido nucleico (i.e., secuencias localizadas en los terminales 5' y 3' del ácido nucleico) en el ADN genómico del organismo de los cuales el ácido nucleico se deriva. Por ejemplo, en varias modalidades, la molécula de ácido nucleico aislada puede contener menos de aproximadamente 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb, o 0.1 kb de secuencias de nucleótidos que naturalmente guardan el flanco de la molécula de ácido nucleico en el ADN genómico de la célula a partir de la cual el ácido nucleico se deriva. Una proteína que

es considerablemente libre de material celular incluye las preparaciones de la proteína que tienen menos de aproximadamente 30%, 20%, 10%, 5%, (en peso seco) de proteína contaminante. Cuando la proteína de la invención o la porción activa biológicamente de estas se produce recombinantemente, preferiblemente el medio de cultivo representa menos de aproximadamente 30%, 20%, 10%, o 5% (en peso seco) de precursores químicos o productos químicos no-proteínas-de-interés.

Los fragmentos y variantes de las secuencias de nucleótidos y proteínas codificadas reveladas por consiguiente también se abarcan por la presente invención. Por “fragmento” se entiende una porción de la secuencia de nucleótidos o una porción de la secuencia de aminoácidos y por lo tanto una proteína codificada así. Los fragmentos de una secuencia de nucleótidos pueden codificar los fragmentos de proteína que retienen la actividad biológica de la P-glicoproteína nativa de la invención y por lo tanto retienen una o más funciones de la P-glicoproteína nativa tal como, por ejemplo, actividad de transporte transmembrana y el enlace ATP. Como alternativa, los fragmentos de una secuencia de nucleótidos que son útiles como sondas de hibridación pueden o no codificar los fragmentos de proteína que retienen la actividad biológica. Así, los fragmentos de una secuencia de nucleótidos pueden fluctuar de al menos aproximadamente 20 nucleótidos, aproximadamente 50 nucleótidos, aproximadamente 100 nucleótidos, y hasta la longitud total de la secuencia de los nucleótidos de la invención.

Un fragmento de una secuencia de nucleótidos del gen de la P-glicoproteína que codifica una porción activa biológicamente de una P-glicoproteína de la invención wiU codifica al menos 70, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 800, 900, 1,000, 1,100, 1,200, 1,300, o 1,400 aminoácidos contiguos, o hasta el número total de aminoácidos presentes en una longitud total P-glicoproteína de la invención (por ejemplo 1,421 aminoácidos para la SEQ ID NO: 9).

Una porción activa biológicamente de una P-glicoproteína de la invención se puede preparar por el aislamiento de una porción de una de las secuencias de nucleótidos del gen de la P-glicoproteína de la invención, expresando la porción codificada de la P-glicoproteína por ejemplo, por expresión recombinante *in vitro*), y evaluando la actividad de la porción de la P-glicoproteína. Las moléculas de ácido nucleico que son fragmentos de un gen de la secuencia de nucleótidos de la P-glicoproteína comprenden al menos 200, 300, 500, 700, 1,000, 1,200, 1,500, 2,000, 2,500, 3,000, 3,500, 4,000, 4,500, 5,000, 5,500, 6,000 nucleótidos, o hasta el número de nucleótidos presentes en una longitud total de la secuencia de nucleótidos de la P-glicoproteína revelada aquí (por ejemplo, 2,139, 1,267, 1,261, 6,827, y 4213 nucleótidos para las SEQ ID NOS: 1-3, y 7-8, respectivamente).

Por “variantes” se entiende secuencias similares considerablemente. Para las secuencias de nucleótidos, las variantes conservadoras incluyen aquellas secuencias que, debido a la degeneración del código genético, codifican la secuencia de aminoácidos de uno de los polipéptidos de la P-glicoproteína de la invención. Las variantes alélicas que ocurren naturalmente tales como estas se pueden identificar con el uso de técnicas de biología molecular bien conocidas, como por ejemplo, con técnicas de reacción en cadena de la polimerasa (PCR) e hibridación como se esquematiza abajo. Las secuencias de nucleótidos variantes también incluyen secuencias de nucleótidos derivadas sintéticamente, tal como aquellas generadas, por ejemplo, utilizando mutagénesis de centros dirigidos pero que aún codifican una proteína de la P-glicoproteína de la invención. Generalmente, las variantes de una secuencia de nucleótidos particular de la invención tendrá al menos aproximadamente 80%, 85%, preferiblemente al menos aproximadamente 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, y más preferiblemente al menos aproximadamente 98%, 99% o más de identidad de la secuencia a aquella secuencia de nucleótidos particular como se determina por los programas de alineamiento de la secuencia descritos aquí en otra parte utilizando parámetros predeterminados.

Por proteína “variante” se entiende una proteína derivada de la proteína nativa por delección (así llamado truncamiento) o adición de uno o más aminoácidos al final del N-terminal y/o C-terminal de la proteína nativa; delección o adición de uno o más aminoácidos en uno o más sitios en la proteína nativa; o sustitución de uno o más aminoácidos en uno o más sitios en la proteína nativa. Las proteínas variantes que se abarcan por la presente invención son activas biológicamente, eso es si continúan teniendo la deseada actividad biológica de la proteína nativa, es decir, actividad de transporte o actividad de enlace ATP según lo descrito aquí. Tales variantes pueden resultar de, por ejemplo, polimorfismo genético o de manipulación humana. Las variantes activas biológicamente de una P-glicoproteína nativa de la invención tendrán al menos aproximadamente 80%, 85%, preferiblemente al menos aproximadamente 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, y más preferiblemente al menos aproximadamente 98%, 99% o más de identidad de la secuencia para la secuencia de aminoácidos para la proteína nativa como se determina por programas de alineamiento de la secuencia descritos en otra parte aquí utilizando parámetros predeterminados. Una variante activa biológicamente de una proteína de la invención puede diferir de esa proteína por solamente 1-15 residuos de aminoácido, solamente 1-10, tal como 6-10, solamente 5, solamente 4, 3, 2, o aún 1, residuo de aminoácido.

Las proteínas de la invención se pueden alterar de varias maneras incluyendo sustituciones, delecciones, truncamientos, e inserciones de aminoácidos. Los métodos para tales manipulaciones generalmente se conocen en el oficio. Por ejemplo, las variantes de secuencias de aminoácidos de las P-glicoproteínas se pueden preparar por mutaciones en el ADN. Los métodos para mutagenesis y alteraciones de secuencias de nucleótidos son bien conocidos en el oficio. Ver por ejemplo, Kunkel (1985) Proc. Natl. Acad. Sci. USA 82:488-492; Kunkel *et al.* (1987) Methods in Enzymol. 154:367-382; US Patent No. 4,873,192; Walker y Gaastra, eds. (1983) Techniques in Molecular Biology (MacMillan Publishing Company, New York) y las referencias citadas allí. La guía en cuanto a las sustituciones de aminoácido apropiadas que no afectan la actividad biológica de la proteína de interés se pueden encontrar en el modelo de Dayhoff *et al.* (1978) Atlas of Protein Sequence y Structure (Natl. Biomed. Res. Found., Washington, D.C.).

Se pueden preferir, sustituciones conservadoras, tal como intercambio de un aminoácido con otro que tiene propiedades similares.

Así, los genes y secuencias de nucleótidos de la invención incluyen tanto las secuencias que ocurren naturalmente así como también las formas mutantes. De igual modo, las proteínas de la invención abarcan tanto las proteínas que ocurren naturalmente así como también las variaciones y formas modificadas de estas. Tales variantes continuarán teniendo la deseada actividad de transporte. Obviamente, las mutaciones que se harán en el ADN que codifica la variante no deben poner la secuencia fuera del marco de lectura y preferiblemente no crearán regiones complementarias que podrían producir una estructura de mRNA secundaria. Ver, EP Patent Application Publication No. 75,444.

No se espera que las delecciones, inserciones, y sustituciones de las secuencias de la proteína abarcadas aquí, produzcan cambios radicales en las características de la proteína. Sin embargo, cuando es difícil predecir el efecto exacto de la sustitución, delección, o inserción anticipada de ser así, alguien de habilidad en el oficio apreciará que el efecto será evaluado por ensayos de detección rutinarios.

Las secuencias de nucleótidos y proteínas variantes también abarcan secuencias de nucleótidos y proteínas derivadas de un procedimiento mutagénico y recombinogénico tal como barajar ADN. Con tal procedimiento, una o más secuencias que codifican diferentes P-glicoproteínas se pueden manipular para crear una secuencia de nucleótidos variante que codifica una P-glicoproteína variante que posee las propiedades deseadas. De esta manera, se generan bibliotecas de polinucleótidos recombinantes a partir de una población de polinucleótidos de secuencia relacionada que comprende las regiones de secuencia que tienen sustancial identidad de la secuencia y se pueden recombinar homológamente *in vitro* o *in vivo*. Por ejemplo, utilizando esta metodología, las fracciones de la secuencia que codifican un dominio de interés se puede barajar entre el gen de la P-glicoproteína de la invención y otros genes conocidos de la P-glicoproteína para obtener un nuevo gen codificante para una proteína con unas mejoradas propiedades de interés, tal como un incremento Km en el caso de una enzima. Las técnicas para barajar ADN son conocidas en el oficio. Ver, por ejemplo, Stemmer (1994) Proc. Natl. Acad. Sci. USA 91:10747-10751; Stemmer (1994) Nature 370:389-391; Cramer *et al.* (1997) Nature Biotech. 15:436-438; Moore *et al.* (1997) J. Mol. Biol. 272:336-347; Zhang *et al.* (1997) Proc. Natl. Acad. Sci. USA 94:4504-4509; Cramer *et al.* (1998) Nature 391:288-291; y U.S. Patent Nos. 5,605,793 y 5,837,458.

Las secuencias de nucleótidos de la invención se pueden utilizar para aislar las secuencias correspondientes de otros organismos, particularmente otras plantas, más particularmente otros monocots. De esta manera, los métodos tales como PCR, hibridación, y similares se pueden utilizar para identificar tales secuencias basadas en su homología de secuencia a las secuencias publicadas aquí. Las secuencias aisladas basadas en su identidad de la secuencia a las secuencias completas publicadas aquí o a los fragmentos de estas se abarcan por la presente invención. Tales secuencias incluyen secuencias que son ortólogos de las secuencias reveladas. Por "ortólogos" se entiende genes derivados de un gen ancestral común y que se encuentran en diferentes especies como un resultado de especiación. Los genes encontrados en diferentes especies se consideran ortólogos cuando sus secuencias de nucleótidos y/o sus secuencias de la proteína codificadas comparten identidad sustancial según lo definido en otra parte aquí. Las funciones de ortólogos son con frecuencia altamente conservadas entre las especies.

En una metodología PCR, los cebadores del desoxirribonucleico se pueden diseñar para utilizar en reacciones PCR para amplificar secuencias de ADN correspondientes a partir del extracto de cADN o ADN genómico de cualquier organismo de interés. Los métodos para diseñar los cebadores PCR y PCR clonado generalmente se conocen en el oficio y se revelan en Sambrook *et al.* (1989) Molecular Cloning: A Laboratory Manual (2nd ed., Cold Spring Harbor Laboratory Press, Plainview, New York). Ver también Innis *et al.*, eds. (1990) PCR Protocols: A Guide to Methods and Applications (Academic Press, New York); Innis and Gelfand, eds. (1995) PCR Strategies (Academic Press, New York); and Innis and Gelfand, eds. (1999) PCR Methods Manual (Academic Press, New York). Métodos conocidos de PCR incluyen, pero no se limitan a, métodos utilizando cebadores apareados, cebadores anidados, cebadores específicos únicos, cebadores degenerados, cebadores de gen específico, cebadores de vector específico, cebadores parcialmente mal apareados, y similares.

En técnicas de hibridación, toda o parte de una secuencia de nucleótidos conocida se utiliza como una sonda que selectivamente hibridiza a otras secuencias correspondientes de nucleótidos presentes en una población de fragmentos de ADN genómico clonado o fragmentos de cADN (i.e., bibliotecas genómicas o de cADN) a partir de un organismo seleccionado. Las sondas de hibridación pueden ser fragmentos de ADN genómico, fragmentos de cADN, fragmentos de ARN, u otros oligonucleótidos, y se pueden marcar con un grupo detectable tal como ³²P, o cualquier otro marcador detectable. Así, por ejemplo, se pueden hacer sondas para hibridación, por marcación de oligonucleótidos sintéticos basadas en el gen de la secuencias de la P-glicoproteína de los nucleótidos de la invención. Los métodos para la preparación de sondas para la hibridación y para la construcción de cADN y bibliotecas genómicas generalmente se conocen en el oficio y se revelan en Sambrook *et al.* (1989) Molecular Cloning: A Laboratory Manual (2nd ed., Cold Spring Harbor Laboratory Press, Plainview, New York).

Por ejemplo, la secuencia *Dw3* completa revelada aquí, o una o más porciones de esta, se pueden utilizar como una sonda capaz de hibridizar específicamente al gen correspondiente de la secuencias de la P-glicoproteína y ARNs mensajero. Para lograr hibridación específica bajo una variedad de condiciones, tales sondas incluyen secuencias que son únicas entre secuencias del gen de la glicoproteína y preferiblemente tienen al menos aproximadamente 10 nucleótidos de longitud, y más preferiblemente al menos aproximadamente 20 nucleótidos de longitud. Tales sondas se

pueden utilizar para amplificar el gen correspondiente de las secuencias de la P-glicoproteína a partir de una planta seleccionada por PCR. Esta técnica se puede utilizar para aislar secuencias codificadas adicionales de una planta deseada o como un ensayo de diagnóstico para determinar la presencia de secuencias codificadas en una planta. Las técnicas de hibridación incluyen detección por hibridación de bibliotecas de ADN colocadas en placas (tanto placas como colonias; ver, por ejemplo, Sambrook *et al.* (1989) *Molecular Cloning: A Laboratory Manual* (2nd ed., Cold Spring Harbor Laboratory Press, Plainview, New York).

La hibridación de tales secuencias se puede realizar bajo condiciones rigurosas. Por “condiciones rigurosas” o “condiciones de hibridación estrictas” se entiende condiciones bajo las cuales una sonda hibridizará a su secuencia diana a un grado de forma detectable mayor que a otras secuencias (por ejemplo, al menos dos veces sobre fondo). Las condiciones estrictas son secuencias-dependientes y serán diferentes en circunstancias diferentes. Controlando la severidad de las condiciones de hibridación y/o lavado, las secuencias diana que son 100% complementarias a la sonda se pueden identificar (homologa por sonda). Como alternativa, las condiciones de severidad se pueden ajustar para permitir algunos apareamientos erróneos en secuencias así que se detectan grados más bajos de similitud (heterólogo por sonda). Generalmente, una sonda tiene menos de aproximadamente 1000 nucleótidos de longitud, preferiblemente menos de 500 nucleótidos de longitud.

Típicamente, las condiciones rigurosas serán aquellas en las cuales la concentración de la sal es menor de aproximadamente 1.5 M de ion de Na, típicamente aproximadamente 0.01 a 1.0 M concentración de ion de Na (u otras sales) a pH 7.0 a 8.3 y la temperatura es al menos aproximadamente 30°C para sondas cortas (por ejemplo, 10 a 50 nucleótidos) y al menos aproximadamente 60°C para sondas largas (por ejemplo, mayor de 50 nucleótidos). Las condiciones estrictas también se pueden lograr con la adición de agentes desestabilizantes tal como formamida. Ejemplares condiciones bajas de severidad incluyen hibridación con un solución reguladora de 30 a 35% de formamida, NaCl 1 M, 1% SDS (dodecil sulfato de sodio) a 37°C, y un lavado en 1X a 2X SSC (20X SSC = 3.0 M NaCl/0.3 M citrato de trisodio) entre 50 a 55°C. Ejemplares condiciones moderadas de severidad incluyen hibridación en 40 a 45% de formamida, NaCl 1 M, 1% SDS a 37°C, y un lavado en 0.5X a 1X SSC entre 55 a 60°C. Ejemplares condiciones altas de severidad incluyen hibridación en 50% de formamida, NaCl 1 M, 1% SDS a 37°C, y un lavado en 0.1X SSC a 60 a 65°C. La duración de la hibridación es generalmente menor de aproximadamente 24 horas, usualmente aproximadamente 4 a aproximadamente 12 horas.

La especificidad es típicamente función de lavados post-hibridación, los factores críticos que son la fuerza iónica y temperatura de la solución final de lavado. Para los híbridos ADN-ADN, la T_m se puede aproximar de la ecuación de Meinkoth y Wahl (1984) *Anal. Biochem.* 138:267-284: $T_m = 81.5^\circ\text{C} + 16.6 (\log M) + 0.41 (\%GC) - 0.61 (\% \text{form}) - 500/L$; donde M es la molaridad de cationes monovalentes, %GC es el porcentaje de nucleótidos de guanosina y citosina en el ADN, % form es el porcentaje de formamida en la solución de hibridación, y L es la longitud del híbrido en pares de base. La T_m es la temperatura (definida bajo fuerza iónica y pH) a la cual el 50% de una secuencia diana complementaria hibridiza a una sonda apareada perfectamente. La T_m se reduce por aproximadamente 1°C para cada 1% de apareamiento erróneo; de esta manera, la T_m , la hibridación, y/o las condiciones de lavado se pueden ajustar para hibridizar a las secuencias de la identidad deseada. Por ejemplo, si se buscan, secuencias con >90% de identidad, la T_m se puede disminuir 10°C. Generalmente, las condiciones rigurosas se seleccionan para ser aproximadamente 5°C menos del punto de fusión térmico (T_m) para la secuencia específica y su complemento a una fuerza iónica definida y pH. Sin embargo, las condiciones severamente rigurosas pueden utilizar una hibridación y/o lavado a 1, 2, 3, o 4°C menor del punto de fusión térmico (T_m); las condiciones moderadamente rigurosas pueden utilizar una hibridación y/o lavado a 6, 7, 8, 9, o 10°C menor del punto de fusión térmico (T_m); condiciones bajas de severidad pueden utilizar una hibridación y/o lavado a 11, 12, 13, 14, 15, o 20°C menor del punto de fusión térmico (T_m). Utilizando la ecuación, la hibridación y composiciones de lavado, y T_m deseados, aquellos de ordinaria habilidad entenderán que las variaciones en la severidad de hibridación y/o soluciones de lavado se describen inherentemente. Si el deseado grado de apareamiento erróneo resulta en una T_m inferior de 45°C (solución acuosa) o 32°C (solución de formamida), se prefiere un incremento de la concentración de SSC así que se puede utilizar, una temperatura alta. Una guía extensa para la hibridación de ácidos nucleicos se encuentra en Tijssen (1993) *Laboratory Techniques in Biochemistry and Molecular Biology-Hybridization with Nucleic Acid Probes*, Part I, Chapter 2 (Elsevier, New York); and Ausubel *et al.*, eds. (1995) *Current Protocols in Molecular Biology*, Chapter 2 (Greene Publishing and Wiley-Interscience, New York). See Sambrook *et al.* (1989) *Molecular Cloning: A Laboratory Manual* (2nd ed., Cold Spring Harbor Laboratory Press, Plainview, New York).

Así, las secuencias aisladas que codifican las P-glicoproteínas y que hibridizan bajo condiciones rigurosas al gen de la secuencias de la P-glicoproteína reveladas aquí, o a los fragmentos de estas, se abarcan por la presente invención. Tales secuencias serán al menos aproximadamente 80% a 85%, y aún 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más homólogos con las secuencias reveladas. Es decir, la identidad de la secuencia de las secuencias puede fluctuar, compartiendo al menos aproximadamente 80% a 85%, y aún al menos aproximadamente 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más de identidad de la secuencia.

Los siguientes términos se utilizan para describir la relaciones de secuencias entre dos o más ácidos nucleicos o polinucleótidos: (a) “secuencia referencia”, (b) “ventana de comparación”, (c) “identidad de la secuencia”, (d) “porcentaje de identidad de la secuencia”, y (e) “identidad sustancial”.

- (a) Según se usa aquí, “secuencia referencia” es una secuencia definida utilizada como una base para la comparación de la secuencia. Una secuencia referencia puede ser un subconjunto o la totalidad de una secuencia

específica; por ejemplo, como un segmento de un cADN de longitud total o secuencia de gen, o el cADN completo o secuencia de gen.

- (b) Según se usa aquí, “ventana de comparación” hace referencia a un segmento contiguo y específico de una secuencia de polinucleótidos, en donde la secuencia de polinucleótidos en la ventana de comparación puede comprender adiciones o deleciones (i.e., huecos) comparado con la secuencia referencia (que no comprenden adiciones o deleciones) para alineación óptima de las dos secuencias. Generalmente, la ventana de comparación es al menos 20 nucleótidos contiguos de longitud, y opcionalmente puede ser 30, 40, 50, 100, o más larga. Aquellos de habilidad en el oficio entenderán que evitar una alta similaridad a una secuencia referencia debido a la inclusión de huecos en la secuencia de polinucleótidos una sanción por hueco típicamente se introduce y se sustrae del número de correspondencias. Los métodos de alineación de secuencias por comparación son bien conocidos en el oficio. Así, la determinación del porcentaje de identidad entre, cualquiera de las dos secuencias se puede lograr utilizando un algoritmo matemático. Ejemplos no-limitantes de tal algoritmo matemático son algorithm of Myers and Miller (1988) CABIOS 4:11-17; the local homology algorithm of Smith *et al.* (1981) Adv. Appl. Math. 2:482; the homology alignment algorithm of Needleman and Wunsch (1970) J. Mol. Biol. 48:443-453; the search-for-similarity method of Pearson and Lipman (1988) Proc. Natl. Acad. Sci. 85:2444-2448; the algorithm of Karlin and Altschul (1990) Proc. Natl. Acad. Sci. USA 87:2264, modified as in Karlin and Altschul (1993) Proc. Natl. Acad. Sci. USA 90:5873-5877.

Implementaciones de ordenador de estos algoritmos matemáticos se pueden utilizar para la comparación de secuencias para determinar la identidad de la secuencia. Tales implementaciones incluyen, pero no se limitan a: CLUSTAL en el programa PC/Gene (disponible de Intelligenetics, Mountain View, California); el programa ALIGN (Version 2.0) y GAP, BESTFIT, BLAST, FASTA, y TFASTA en el Wisconsin Genetics Software Package, Version 8 (disponible de Genetics Computer Group (GCG), 575 Science Drive, Madison, Wisconsin, USA). Las alineaciones utilizando estos programas se pueden realizar utilizando los parámetros predeterminados. El programa CLUSTAL se describe bien por Higgins *et al.* (1988) Gene 73:237-244 (1988); Higgins *et al.* (1989) CABIOS 5:151-153; Corpet *et al.* (1988) Nucleic Acids Res. 16:10881-90; Huang *et al.* (1992) CABIOS 8:155-65; y Pearson *et al.* (1994) Meth. Mol. Biol. 24:307-331. El programa ALIGN se basa en el algoritmo de Myers and Miller (1988) *supra*. Una tabla de peso del residuo PAM120, una penalización de la longitud del hueco de 12, y una penalización del hueco de 4 se puede utilizar con el programa ALIGN cuando se comparan las secuencias de aminoácidos. Los programas BLAST de Altschul *et al.* (1990) J. Mol. Biol. 215:403 se basan en el algoritmo de Karlin y Altschul (1990) *supra*. Investigaciones de nucleótidos BLAST se pueden realizar con el programa BLASTN, puntuación = 100, longitud de la palabra = 12, para obtener las secuencias de nucleótidos homólogos a una secuencia de nucleótidos que codifica una proteína de la invención. Las investigaciones de proteínas BLAST se pueden realizar con el programa BLASTX, puntuación = 50, longitud de la palabra = 3, para obtener secuencias de aminoácidos homólogos a una proteína o polipéptido de la invención. Para obtener alineaciones discontinuas para propósitos de comparación, Gapped BLAST (en BLAST 2.0) puede ser utilizado según lo descrito en Altschul *et al.* (1997) Nucleic Acids Res. 25:3389. Como alternativa, PSI-BLAST (en BLAST 2.0) se pueden utilizar para realizar una investigación repetida que detecta las relaciones distantes entre moléculas. Ver Altschul *et al.* (1997) *supra*. Cuando se utiliza BLAST, Gapped BLAST, PSIBLAST, se pueden utilizar, los parámetros predeterminados de los programas respectivos (por ejemplo, BLASTN para secuencias de nucleótidos, BLASTX para proteínas). Ver <http://www.ncbi.nlm.nih.gov>. La alineación también se puede realizar manualmente por inspección. A menos que se indique de otra manera, los valores de identidad de la secuencia/similaridad proporcionados aquí se refieren al valor obtenido utilizando la Version 10 GAP utilizando los siguientes parámetros: % de identidad utilizando GAP Peso de 50 y Longitud de Peso de 3; % similaridad utilizando Gap Peso de 12 y Longitud de Peso de 4, o cualquier programa equivalente. Por “programa equivalente” se entiende cualquier comparación del programa de la secuencia que, para cualquiera de las dos secuencias en cuestión, genera una alineación que tiene equivalencias de nucleótido idéntico o residuo de aminoácido y un porcentaje de identidad idéntico de la secuencia cuando se compara a la correspondiente alineación generada por el programa preferido.

GAP usa el algoritmo de Needleman and Wunsch (1970) J. Mol. Biol. 48: 443-453, para encontrar la alineación de dos secuencias completas que maximiza el número de equivalencias y minimiza el número de huecos. GAP considera todas las posibles alineaciones y posiciones del hueco y crea la alineación con el mayor número de bases apareadas y los pocos huecos. Esto permite el suministro de una penalización de una creación del hueco y una penalización de la extensión del hueco en unidades de bases apareadas. GAP debe hacer un beneficio de la penalización de una creación del número de huecos de equivalencias para cada hueco que se inserta. Si se selecciona, una penalización de la extensión del hueco mayor de cero, GAP debe, además, hacer un beneficio para cada hueco insertado de la longitud de las veces del hueco la penalización de la extensión del hueco. La penalización predeterminada de una creación de los valores del hueco y la penalización de la extensión del hueco en la Version 10 del Wisconsin Genetics Software Package para las secuencias de la proteína son 8 y 2, respectivamente. Para las secuencias de nucleótidos la penalización predeterminada de una creación del hueco es 50 mientras que la penalización predeterminada de la extensión del hueco es 3. Las penalizaciones de creación del hueco y extensión del hueco se pueden expresar como un número entero seleccionado del grupo de números enteros que consisten desde 0 a 200. Así, por ejemplo, las penalizaciones de creación del hueco y extensión del hueco pueden ser 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65 o más grandes.

GAP presenta un miembro de la familia de mejores alineaciones. Puede haber muchos miembros de esta familia, pero ningún otro miembro tiene una mejor calidad. GAP muestra cuatro figuras de mérito para las alineaciones:

Calidad, Índice, Identidad, y Similaridad. La Calidad es la métrica maximizada con el fin de alinear las secuencias. El Índice es la calidad dividida por el número de bases en el segmento más corto. El porcentaje de Identidad es el porcentaje de los símbolos que realmente corresponde. El porcentaje de Similaridad es el porcentaje de los símbolos que son similares. Los símbolos que están a través de los huecos se ignoran. Una similaridad se puntúa cuando el valor matriz de puntuación para un par de símbolos es mayor de o igual a 0.50, el umbral similaridad. La matriz de puntuación utilizada en la Version 10 del Wisconsin Genetics Software Package es BLOSUM62 (ver Henikoff y Henikoff (1989) Proc. Natl. Acad. Sci. USA 89:10915).

(c) Según se usa aquí, “identidad de la secuencia” o “identidad” en el contexto de dos secuencias de ácido nucleico o polipéptido hace referencia a los residuos en las dos secuencias que son las mismas cuando se alinean para una correspondencia máxima sobre una ventana de comparación específica. Cuando el porcentaje de identidad de la secuencia se utiliza en referencia a las proteínas, se reconoce que las posiciones del residuo que no son idénticas con frecuencia difieren por sustituciones de aminoácido conservadoras, donde los residuos de aminoácidos se sustituyen por otros residuos de aminoácidos con propiedades químicas similares (por ejemplo, carga o hidrofobicidad) y por consiguiente no cambian las propiedades funcionales de la molécula. Cuando las secuencias difieren en sustituciones conservadoras, el porcentaje de identidad de la secuencia se puede ajustar hacia arriba para corregir la naturaleza conservadora de la sustitución. Las secuencias que difieren mediante tales sustituciones conservadoras se dice que tiene “similaridad de secuencia” o “similaridad”. Los medios para hacer este ajuste son bien conocidos por aquellos de habilidad en el oficio. Típicamente esto involucra la puntuación de una sustitución conservadora como parcial antes que una incompatibilidad completa, por consiguiente incrementando el porcentaje de identidad de la secuencia. Así, por ejemplo, donde un aminoácido idéntico, se le da una puntuación de 1 y a una sustitución no-conservadora se le da una puntuación de cero, a una sustitución conservadora se le da una puntuación entre cero y 1. La puntuación de las sustituciones conservadoras se calcula, por ejemplo, según lo implementado en el programa PC/GENE (Intelligenetics, Mountain View, California).

(d) Según se usa aquí, “porcentaje de identidad de la secuencia” significa el valor determinado, comparando dos secuencias alineadas optimamente sobre una ventana de comparación, en donde la porción de la secuencia de polinucleótidos en la ventana de comparación puede comprender adiciones o deleciones (i.e., huecos) según se compara con la secuencia referencia (que no comprende adiciones o deleciones) para una alineación óptima de las dos secuencias. El porcentaje se calcula determinando el número de posiciones a las cuales la base de ácido nucleico idéntico o residuo de aminoácido ocurre en ambas secuencias para producir el número de posiciones apareadas; dividiendo el número de posiciones apareadas por el número total de posiciones en la ventana de comparación, y multiplicando el resultado por 100 para producir el porcentaje de identidad de la secuencia.

(e) (i) El término “identidad sustancial” de las secuencias del polinucleótido significa que un polinucleótido contiene una secuencia que tiene al menos 80% de identidad, más preferiblemente al menos 90%, y más preferiblemente al menos 95%, comparado con una secuencia referencia utilizando uno de los programas de alineación descritos utilizando parámetros estándar. Alguien de habilidad en el oficio reconocerá que estos valores se puede ajustar apropiadamente para determinar la correspondiente identidad de las proteínas codificadas por dos secuencias de nucleótidos tomando en cuenta la degeneración codón, similaridad del aminoácido, posicionamiento del marco de lectura, y similares. La sustancial identidad de las secuencias de aminoácidos para estos propósitos normalmente significa una identidad de la secuencia de al menos 80%, 90%, y más preferiblemente al menos 95%. Otro indicio de que las secuencias de los nucleótidos son idénticas considerablemente es si dos moléculas hibridizan la una con la otra bajo condiciones rigurosas. Generalmente, las condiciones rigurosas se seleccionan para ser aproximadamente 5°C menor del punto de fusión térmico (T_m) para la secuencia específica a una fuerza iónica y pH definidos. Sin embargo, las condiciones rigurosas abarcan temperaturas en el rango de aproximadamente 1°C a aproximadamente 20°C menor del T_m , dependiendo del grado deseado de severidad según lo calificado aquí por otra parte. Los ácidos nucleicos que no hibridizan cada uno con el otro bajo condiciones rigurosas todavía son considerablemente idénticos, si los polipéptidos que codifican son considerablemente idénticos. Esto puede ocurrir, por ejemplo, cuando una copia de un ácido nucleico se crea utilizando la máxima degeneración del codón permitida por el código genético. Un indicio de que las dos secuencias de ácido nucleico son considerablemente idénticas es cuando el polipéptido codificado por el primer ácido nucleico es inmunológicamente reactivo en cruz con el polipéptido codificado por el segundo ácido nucleico.

(e) (ii) El término “identidad sustancial” en el contexto de un péptido indica que un péptido comprende una secuencia con al menos 80% de identidad de la secuencia a una secuencia referencia, más preferiblemente 85%, más preferiblemente al menos 90% o 95% de identidad de la secuencia para la secuencia referencia sobre una ventana de comparación específica. Preferiblemente, la alineación óptima se conduce utilizando la homología del algoritmo de alineación de Needleman *et al.* (1970) J. Mol. Biol. 48:443. Un indicio de que las dos secuencias del péptido son considerablemente idénticas es que un péptido es inmunológicamente reactivo con los anticuerpos contruidos contra el segundo péptido. Así, un péptido es considerablemente idéntico a un segundo péptido, por ejemplo, donde los dos péptidos difieren solo por una sustitución conservadora. Los péptidos que son “considerablemente similares” comparten secuencias como se indica arriba excepto que las posiciones del residuo que no son idénticas pueden diferir por cambios conservadores del aminoácido.

El uso del término “construcciones del nucleótido” aquí no tiene la intención de limitar la presente invención a construcciones del nucleótido que contienen ADN. Aquellos de ordinaria habilidad en el oficio reconocerán que las construcciones del nucleótido, particularmente polinucleótidos y oligonucleótidos, se comprenden de ribonucleótidos y combinaciones de ribonucleótidos y desoxiribonucleótidos también se pueden emplear en los métodos revelados aquí. Así, las construcciones del nucleótido de la presente invención abarcan todas las construcciones del nucleótido que se pueden emplear en los métodos de la presente invención para convertir las plantas incluyendo, pero no limitando a, aquellos comprendidos de desoxiribonucleótidos, ribonucleótidos, y combinaciones de estos. Tales desoxiribonucleótidos y ribonucleótidos incluyen ambos, moléculas que ocurren naturalmente y análogos sintéticos. Las construcciones del nucleótido de la invención también abarcan todas las formas de construcciones del nucleótido incluyendo, pero no limitando a, formas de cadena sencilla, formas de doble cadena, horquilla, estructuras de bucle-y-tallo, y similares.

Se puede emplear una construcción del nucleótido es decir capaz de dirigir, en una planta transformada, la expresión de al menos una proteína, o al menos un ARN, tal como, por ejemplo, un ARN antisentido que es complementario a por los menos una porción de un mRNA. Típicamente tal construcción del nucleótido se comprende de una secuencia codificante para una proteína o un ARN ligado operablemente a las regiones 5' y 3' reguladoras transcripcionales. Como alternativa, también se reconoce que los métodos de la invención pueden emplear una construcción del nucleótido que no sea capaz de dirigir, en una planta transformada, la expresión de una proteína o un ARN.

Además, se reconoce que los métodos de la presente invención no dependen de la incorporación en el genoma de la construcción del nucleótido total que comprende una secuencia de nucleótido de la P-glicoproteína, solo que la planta o célula de esta se altera como un resultado de la introducción de la construcción del nucleótido en una célula. En una modalidad de la invención, el genoma se puede alterar siguiendo la introducción de la construcción del nucleótido en una célula. Por ejemplo, la construcción del nucleótido, o cualquier parte de esta, puede incorporar en el genoma de la planta. Las alteraciones al genoma de la presente invención incluyen, pero no se limitan a, adiciones, deleciones, y sustituciones de nucleótidos en el genoma. Mientras que los métodos de la presente invención no dependen de adiciones, deleciones, o sustituciones de ningún número de nucleótidos particular, se reconoce que tales adiciones, deleciones, o sustituciones comprenden al menos un nucleótido.

Las construcciones del nucleótido de la invención también abarcan construcciones del nucleótido que se pueden emplear en métodos alterando o mutando una secuencia de nucleótidos genómica en un organismo, incluyendo, pero no limitando a, vectores quiméricos, vectores mutacionales quiméricos, vectores de reparación quiméricos, oligonucleótidos mezclados-dobles, oligonucleótidos quiméricos auto-complementarios, y oligonucleobases recombinogénicas. Tales construcciones del nucleótido y los métodos de uso, tal como, por ejemplo, quimeraplastia, son conocidos en el oficio. La quimeraplastia involucra el uso de dichas construcciones del nucleótido para introducir cambios de sitio-específico en la secuencia de ADN genómico dentro de un organismo. Ver, Patente U.S. Nos. 5,565,350; 5,731,181; 5,756,325; 5,760,012; 5,795,972; y 5,871,984. Ver también, WO 98/49350, WO 99/07865, WO 99/25821, y Beetham *et al.* (1999) Proc. Natl. Acad. Sci. USA 96:8774-8778.

La invención abarca el uso de métodos, tal como, por ejemplo, quimeraplastia para alterar los genes de la P-glicoproteína en las plantas. Tales alteraciones incluyen, por ejemplo, cambios en la secuencia codificante que altera la secuencia de aminoácidos de the P-glicoproteína codificada por consiguiente, resultando en una reducción en, o pérdida de, la función de la P-glicoproteína codificada por ese gen.

Las secuencias de nucleótidos de la P-glicoproteína de la invención se proporcionan en casetes de expresión para la expresión en la planta de interés. El casete incluirá las secuencias reguladoras 5'- y 3' ligadas operablemente a una secuencia de nucleótidos de la P-glicoproteína de la invención. Por “ligado operablemente” se entiende un enlace funcional entre un promotor y una segunda secuencia, en donde la secuencia promotor inicia y media la transcripción de la secuencia de ADN correspondiente a la segunda secuencia. Generalmente, ligado operablemente significa que las secuencias de ácido nucleico que se unen están contiguas y, cuando sea necesario unir dos regiones que codifican la proteína, contiguas y en el mismo marco de lectura. El casete puede adicionalmente contener al menos un gen adicional que se cotransforma en el organismo. Como alternativa, el o los genes adicionales pueden ser proporcionados en casetes de expresión múltiples.

Tal casete de expresión se suple con una pluralidad de sitios de restricción para la inserción de la secuencia de nucleótidos de la P-glicoproteína para estar bajo la regulación transcripcional de las regiones reguladoras. El casete de expresión puede adicionalmente contener genes marcadores genéticos.

El casete de expresión incluirá en la dirección 5'-3' de la transcripción, una región de iniciación transcripcional y translacional, una secuencia de nucleótidos de la P-glicoproteína de la invención, y una región de terminación transcripcional y translacional funcional en las plantas. La región de iniciación transcripcional, el promotor, puede ser nativa o análoga o extraña o heteróloga a la planta huésped. Adicionalmente, el promotor puede tener la secuencia natural o alternativamente una secuencia sintética. Por “extraña” se entiende que la región de iniciación transcripcional no se encuentra en la planta nativa en la cual se introduce la región de iniciación transcripcional.

Mientras que esta puede ser preferible para expresar las secuencias utilizando promotores heterólogos, los promotores de la secuencia nativa se pueden utilizar. Tales construcciones cambiarían los niveles de expresión de una P-glicoproteína en la planta o célula de la planta. Así, el fenotipo de la planta o célula de la planta se altera.

La región de terminación puede ser nativa con la región de iniciación transcripcional, puede ser nativa con la secuencia de ADN ligada operablemente de interés, o puede ser derivada de otra fuente. Las regiones de terminación convenientes están disponibles del Ti-plásmid de *A. tumefaciens*, tal como las regiones de terminación octopina sintasa y nopalina sintasa. Ver también Guerineau *et al.* (1991) Mol. Gen. Genet. 262:141-144; Proudfoot (1991) Cell 64:671-674; Sanfacon *et al.* (1991) Genes Dev. 5:141-149; Mogen *et al.* (1990) Plant Cell 2:1261-1272; Munroe *et al.* (1990) Gene 91:151-158; Ballas *et al.* (1989) Nucleic Acids Res. 17:7891-7903; and Joshi *et al.* (1987) Nucleic Acid Res. 15:9627-9639.

Cuando es apropiado, el o los genes se pueden optimizar para incrementar la expresión en la planta transformada. Es decir, los genes se pueden sintetizar utilizando codones de planta-preferidos para expresión mejorada. Los métodos están disponibles en el oficio para sintetizar los genes de planta-preferidos. Ver, por ejemplo, Patente U.S. Nos. 5,380,831, y 5,436,391, and Murray *et al.* (1989) Nucleic Acids Res. 17:477-498.

Las modificaciones de secuencia adicional son conocidas para incrementar la expresión del gen en un huésped celular. Estos incluyen la eliminación de secuencias que codifican señales de poliadenilación falsas, señales de sitio de empalme exón-intrón, repetición como transposón, y otras secuencias bien caracterizadas que pueden ser nocivas para la expresión del gen. El contenido G-C de la secuencia se puede ajustar a los niveles promedio para un huésped celular dado, según se calcula por referencia a los genes conocidos expresados en la célula huésped. Cuando es posible, la secuencia se modifica para evitar las estructuras de mRNA secundaria de horquilla predichas.

Los casetes de expresión pueden adicionalmente contener secuencias líder-5' en la construcción del casete de expresión. Tales secuencias líderes pueden actuar para incrementar la translación. Los líderes de translación se conocen en el oficio e incluyen: líderes picornavirus, por ejemplo, EMC leader (Encephalomyocarditis 5'-noncoding region) (Elroy-Stein *et al.* (1989) PNAS USA 86:6126-6130); potyvirus leaders, por ejemplo, TEV leader (Tobacco Etch Virus) (Allison *et al.* (1986); MDMV leader (Maize Dwarf Mosaic Virus); Virology 154:9-20), and human immunoglobulin heavy-chain binding protein (BiP), (Macejak *et al.* (1991) Nature 353:90-94); untranslated leader from the coat protein mRNA of alfalfa mosaic virus (AMV RNA 4) (Jobling *et al.* (1987) Nature 325:622-625); tobacco mosaic virus leader (TMV) (Gallie *et al.* (1989) in Molecular Biology of RNA, ed. Cech (Liss, New York), pp. 237-256); and maize chlorotic mottle virus leader (MCMV) (Lommel *et al.* (1991) Virology 81:382-385). Ver también, Della-Cioppa *et al.* (1987) Plant Physiol. 84:965-968. Otros métodos conocidos para incrementar la translación también se pueden utilizar, por ejemplo, intrones, y similares.

En la preparación del casete de expresión, se pueden manipular diversos fragmentos de ADN, para prever las secuencias de ADN en la orientación correcta y, como apropiadas, en el marco de lectura correcto. Hacia este fin, los adaptadores o conectores se pueden emplear para unir los fragmentos de ADN u otras manipulaciones se pueden involucrar para prever convenientes sitios de restricción, eliminación de ADN innecesario, eliminación de sitios de restricción, o similares. Para este propósito, se pueden involucrar, mutagénesis *in vitro*, reparación del cebador, restricción, fijación, resustituciones, por ejemplo, transiciones y transversiones.

Se reconoce que con las secuencias de nucleótidos de la invención, las construcciones antisentido, complementarias para al menos una porción del ARN mensajero (mARN) para el gen de la secuencias de la P-glicoproteína se pueden construir. Los nucleótidos antisentido se construyen para hibridizar con el mARN correspondiente. Las modificaciones de las secuencias antisentido se pueden hacer tan largas como las secuencias que hibridizan a, e interfieren con expresión del mARN correspondiente. De esta manera, se pueden utilizar, las construcciones antisentido que tienen 70%, preferiblemente 80%, más preferiblemente 85% de identidad de la secuencia para las secuencias diana correspondientes. Adicionalmente, las porciones de los nucleótidos antisentido se pueden utilizar para desestabilizar la expresión del gen diana. Generalmente, se pueden utilizar, las secuencias de al menos 50 nucleótidos, 100 nucleótidos, 200 nucleótidos, o más grandes.

Las secuencias de nucleótidos de la presente invención también se pueden utilizar en la orientación de sentido para suprimir la expresión de genes endógenos en las plantas. Los métodos para suprimir la expresión del gen en las plantas utilizando secuencias de nucleótidos en la orientación de sentido, también conocidos como métodos de cosupresión, se conocen en el oficio. Los métodos generalmente involucran convertir las plantas con una construcción del nucleótido que comprende un promotor que conduce la expresión en una planta ligada operablemente a al menos una porción de una secuencia de nucleótidos que corresponde a la transcripción del gen endógeno. Típicamente, tal secuencia de nucleótidos tiene identidad sustancial de la secuencia con la secuencia de la transcripción del gen endógeno, preferiblemente mayor de aproximadamente 65% de identidad de la secuencia, más preferiblemente mayor de aproximadamente 85% de identidad de la secuencia, más preferiblemente mayor de aproximadamente 95% de identidad de la secuencia. Ver, Patente U.S. Nos. 5,283,184 y 5,034,323.

Generalmente, el casete de expresión comprenderá un gen marcador genético para la selección de células transformadas. Los genes marcadores genéticos se utilizan para la selección de células o tejidos transformados. Los genes marcadores incluyen genes que codifican la resistencia antibiótica, tal como aquellos que codifican neomicina fosfotransferasa II (NEO) e higromicina fosfotransferasa (HPT), así como también los genes que confieren resistencia a los compuestos herbicidas, tal como glufosinato de amonio, bromoxinil, imidazolinonas, y 2,4-diclorofenoxiacetato (2,4-D). Ver en general, Yarranton (1992) Curr. Opin. Biotech. 3:506-511; Christopherson *et al.* (1992) Proc. Natl. Acad. Sci. USA 89:6314-6318; Yao *et al.* (1992) Cell 71:63-72; Reznikoff (1992) Mol. Microbiol. 6:2419-2422; Barkley *et al.* (1980) in The Operon, pp. 177-220; Hu *et al.* (1987) Cell 48:555-566; Brown *et al.* (1987) Cell 49:603-612; Figge

- el al. (1988) *Cell* 52:713-722; Deuschle *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86:5400-5404; Fuerst *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86:2549-2553; Deuschle *et al.* (1990) *Science* 248:480-483; Gossen (1993) Ph.D. Thesis, University of Heidelberg; Reines *et al.* (1993) *Proc. Natl. Acad. Sci. USA* 90:1917-1921; Labow *et al.* (1990) *Mol. Cell. Biol.* 10:3343-3356; Zambretti *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:3952-3956; Baim *et al.* (1991) *Proc. Natl. Acad. Sci. USA* 88:5072-5076; Wyborski *et al.* (1991) *Nucleic Acids Res.* 19:4647-4653; Hillenand-Wissman (1989) *Topics Mol. Struc. Biol.* 10:143-162; Degenkolb *et al.* (1991) *Antimicrob. Agents Chemother.* 35:1591-1595; Kleinschmidt *et al.* (1988) *Biochemistry* 27:1094-1104; Bonin (1993) Ph.D. Thesis, University of Heidelberg; Gossen *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:5547-5551; Oliva *et al.* (1992) *Antimicrob. Agents Chemother.* 36:913-919; Hlavka *et al.* (1985) *Handbook of Experimental Pharmacology*, Vol. 78 (Springer-Verlag, Berlin); Gill *et al.* (1988) *Nature* 334:721-724.

La anterior lista de genes marcadores genéticos no tiene la intención de ser limitante. Cualquier gen marcador genético se puede utilizar en la presente invención.

- Un número de promotores se puede utilizar en la práctica de la invención. Los promotores se pueden seleccionar basándose en la medición, localización y nivel de expresión deseada de los genes de la P-glicoproteína en una planta. Constitutivo, los promotores de tejido preferido, patógeno-inducible, lesión-inducible y regulables químicamente se pueden utilizar en la práctica de la invención.

- Tales promotores constitutivos incluyen, por ejemplo, el promotor del núcleo del promotor Rsyn7 y otros promotores constitutivos revelados en WO 99/43838 y Patente U.S. No. 6,072,050; el promotor núcleo CaMV 35S (Odell *et al.* (1985) *Nature* 313:810-812); arroz actina (McElroy *et al.* (1990) *Plant Cell* 2:163-171); ubiquitin (Christensen *et al.* (1989) *Plant Mol. Biol.* 12:619-632 and Christensen *et al.* (1992) *Plant Mol. Biol.* 18:675-689); pEMU (Last *et al.* (1991) *Theor. Appl. Genet.* 81:581-588); MAS (Velten *et al.* (1984) *EMBO J.* 3:2723-2730); ALS promotor (U.S. Application Serial No. 08/409,297), y similares. Otros promotores constitutivos incluyen, por ejemplo, Patente U.S. Nos. 5,608,149; 5,608,144; 5,604,121; 5,569,597; 5,466,785; 5,399,680; 5,268,463; y 5,608,142.

- Los promotores de tejido-preferido pueden ser utilizados para apuntar a la mejora de la expresión de la P-glicoproteína dentro de un tejido de planta particular. Los promotores de tejido-preferido incluyen Yamamoto *et al.* (1997) *Plant J.* 12(2):255-265; Kawamata *et al.* (1997) *Plant Cell Physiol.* 38(7):792-803; Hansen *et al.* (1997) *Mol. Gen. Genet.* 254(3):337-343; Russell *et al.* (1997) *Transgenic Res.* 6(2):157-168; Rinehart *et al.* (1996) *Plant Physiol.* 112(3):1331-1341; Van Camp *et al.* (1996) *Plant Physiol.* 112 (2):525-535; Canevascini *et al.* (1996) *Plant Physiol.* 112(2):513-524; Yamamoto *et al.* (1994) *Plant Cell Physiol.* 35(5): 773-778; Lam (1994) *Results Probl. Cell Differ.* 20:181-196; Orozco *et al.* (1993) *Plant Mol Biol.* 23(6):1129-1138; Matsuoka *et al.* (1993) *Proc Natl. Acad Sci. USA* 90(20):9586-9590; y Guevara-García *et al.* (1993) *Plant J.* 4(3): 495-505. Tales promotores se pueden modificar, si es necesario, para expresión débil.

- Los promotores de hoja-preferida incluyen, Yamamoto *et al.* (1997) *Plant J.* 12(2):255-265; Kawamata *et al.* (1997) *Plant Cell Physiol.* 38(7):792-803; Hansen *et al.* (1997) *Mol. Gen. Genet.* 254(3):337-343; Russell *et al.* (1997) *Transgenic Res.* 6(2):157-168; Rinehart *et al.* (1996) *Plant Physiol.* 112(3):1331-1341; Van Camp *et al.* (1996) *Plant Physiol.* 112 (2):525-535; Canevascini *et al.* (1996) *Plant Physiol.* 112(2):513-524; Yamamoto *et al.* (1994) *Plant Cell Physiol.* 35(5): 773-778; Lam (1994) *Results Probl. Cell Differ.* 20:181-196; Orozco *et al.* (1993) *Plant Mol. Biol.* 23 (6):1129-1138; Matsuoka *et al.* (1993) *Proc. Natl. Acad. Sci. USA* 90(20):9586-9590; y Guevara-García *et al.* (1993) *Plant J.* 4(3): 495-505.

- Los promotores de raíz-preferidos se conocen y se pueden seleccionar a partir de varios disponibles de la literatura o aislada de novo a partir de diversas especies compatibles. Ver, por ejemplo, Hire *et al.* (1992) *Plant Mol. Biol.* 20(2): 207-218 (soybean root-preferred glutamine synthetase gene); Keller and Baumgartner (1991) *Plant Cell* 3(10):1051-1061 (root-preferred control element in the GRP 1.8 gene of French bean); Sanger *et al.* (1990) *Plant Mol. Biol.* 14 (3):433-443 (root-preferred promoter of the mannopine synthase (MAS) gene of *Agrobacterium tumefaciens*); and Miao *et al.* (1991) *Plant Cell* 3 (1):11-22 (full-length cDNA clone encoding cytosolic glutamine synthetase (GS), which is expressed in roots and root nodules of soybean). Ver también Bogusz *et al.* (1990) *Plant Cell* 2(7):633-641, donde se describen, dos promotores de raíz preferidos aislados de los genes de hemoglobina del nonlegume *Parasponia andersonii* nitrógeno-fijo y el nonlegume *Trema tomentosa* relacionado nitrógeno-no-fijo. Los promotores de estos genes se ligaron a un gen reportero β -glucoronidasa y se introduce en ambos el nonlegume *Nicotiana tabacum* y el legume *Lotus corniculatus*, y en ambos ejemplos la actividad del promotor de raíz-preferida se preservó. Leach y Aoyagi (1991) describen su análisis de los promotores de los genes que inducen-raíz expresan altamente rolC y rolD de *Agrobacterium rhizogenes* (ver *Plant Science* (Limerick) 79(1):69-76). Ellos concluyen que el mejorador y determinantes de ADN de tejido-preferido se disocian en aquellos promotores. Teeri *et al.* (1989) utilizó fusión del gen a lacZ para mostrar que el gen del T-ADN de *Agrobacterium* que codifica la octopina sintasa es especialmente activo en la epidermis de la punta de la raíz y que el gen de TR2' es raíz preferida en la planta intacta y se estimula por lesiones en tejido de hoja, una combinación de características deseables especialmente para utilizar con un gen insecticida o larvicida (ver *EMBO J.* 8(2):343-350). El gen de TR1, combinado a *nptII* (neomicina fosfotransferasa II) mostró características similares. Los promotores adicionales de raíz preferidos incluyen el gen promotor VfENOD-GRP3 (Kuster *et al.* (1995) *Plant Mol. Biol.* 29(4):759-772); y rolB promotor (Capana *et al.* (1994) *Plant Mol. Biol.* 25(4):681-691. Ver también Patente U.S. Nos. 5,837,876; 5,750,386; 5,633,363; 5,459,252; 5,401,836; 5,110,732; y 5,023,179.

Generalmente, será beneficioso expresar el gen de un promotor inducible, particularmente de un promotor patógeno-inducible. Tales promotores incluyen aquellos de las proteínas patogénesis-relacionadas (proteínas PR), que se inducen después de la infección por un patógeno; por ejemplo, proteínas PR, proteínas SAR, beta-1,3-glucanasa, quitinasa, etc. Ver, por ejemplo, Redolfi *et al.* (1983) *Neth. J. Plant Pathol.* 89:245-254; Uknes *et al.* (1992) *Plant Cell* 4:645-656; y Van Loon (1985) *Plant Mol. Virol.* 4:111-116. Ver también the copending applications entitled "Inducible Maize Promoters", U.S. Application Serial No. 60/076,100, filed February 26, 1998, y U.S. Application Serial No. 60/079,648, filed March 27, 1998, ambas de las cuales corresponden a la aplicación PCT WO 99/43819, publicada el septiembre 2 de 1999.

De interés son los promotores que se expresan localmente en o cerca del sitio de infección del patógeno. Ver, por ejemplo, Marineau *et al.* (1987) *Plant Mol. Biol.* 9:335-342; Matton *et al.* (1989) *Molecular Plant-Microbe Interactions* 2:325-331; Somsisch *et al.* (1986) *Proc. Natl. Acad. Sci. USA* 83:2427-2430; Somsisch *et al.* (1988) *Mol. Gen. Genet.* 2:93-98; y Yang (1996) *Proc. Natl. Acad. Sci. USA* 93:14972-14977. Ver también, Chen *et al.* (1996) *Plant J.* 10:955-966; Zhang *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:2507-2511; Warner *et al.* (1993) *Plant J.* 3:191-201; Siebertz *et al.* (1989) *Plant Cell* 1:961-968; Patente U.S. No. 5,750,386 (nematodo-inducible); y las referencias citadas aquí. De particular interés es el promotor inducible para el gen PRms del maíz, cuya expresión se induce por el patógeno *Fusarium moniliforme* (ver, por ejemplo, Cordero *et al.* (1992) *Physiol. Mol. Plant Pathol.* 41:189-200).

Adicionalmente, como patógenos encuentran la entrada en las plantas con heridas o daño del insecto, un promotor herida-inducible se puede utilizar en las construcciones de la invención. Tales promotores herida-inducible incluyen el gen inhibidor de proteinasa de patata (pin II) (Ryan (1990) *Ann. Rev. Phytopath.* 28:425-449; Duan *et al.* (1996) *Nature Biotechnology* 14:494-498); *wun1* y *wun2*, US Patent No. 5,428,148; *win1* y *win2* (Stanford *et al.* (1989) *Mol. Gen. Genet.* 215:200-208); *systemin* (McGurl *et al.* (1992) *Science* 225:1570-1573); *WIP1* (Rohmeier *et al.* (1993) *Plant Mol. Biol.* 22:783-792; Eckelkamp *et al.* (1993) *FEBS Letters* 323:73-76); gen *MPI* (Corderok *et al.* (1994) *Plant J* 6(2):141-150).

Los promotores regulados químicamente se pueden utilizar para modular la expresión de un gen en una planta con la aplicación de un regulador químico exógeno. Dependiendo del objetivo, el promotor puede ser un promotor químico-inducible, donde la aplicación del químico induce la expresión del gen, o un promotor químico-reprimible, donde la aplicación del químico reprime la expresión del gen. Los promotores químicamente inducibles se conocen en el oficio e incluyen, pero no se limitan a, el promotor *ln2-2* del maíz, el cual se activa por protectores herbicidas bencenosulfonamida, el promotor *GST* del maíz, el cual se activa por compuestos electrofílicos hidrofóbicos que se utilizan como herbicidas pre-emergentes, y el promotor *PR-1a* del tabaco, que se activa por ácido salicílico. Otros promotores regulados químicamente de interés incluyen promotores esteroide-sensible (ver, por ejemplo, el promotor glucocorticoide-inducible en Schena *et al.* (1991) *Proc. Natl. Acad. Sci. USA* 88:10421-10425 and McNellis *et al.* (1998) *Plant J* 14(2):247-257) y los promotores tetraciclina-inducible y tetraciclina-reprimible (ver, por ejemplo, Gatz *et al.* (1991) *Mol. Gen. Genet.* 227:229-237, y Patente U.S. Nos. 5,814,618 y 5,789,156).

Los protocolos de transformación así como también los protocolos para introducir las secuencias de nucleótidos en las plantas pueden variar dependiendo del tipo de planta o célula de la planta, i.e., monocotiledónea o dicotiledónea, dirigiendo para la transformación. Los métodos apropiados para introducir las secuencias de nucleótidos en las células de las plantas y subsiguiente inserción en el genoma de la planta incluyen microinyección (Crossway *et al.* (1986) *Biotechniques* 4:320-334), electroporación (Riggs *et al.* (1986) *Proc. Natl. Acad. Sci. USA* 83:5602-5606, transformación *Agrobacterium*-mediada (Townsend *et al.*, Patente U.S. No. 5,563,055; Zhao *et al.*, Patente U.S. No. 5,981,840), transferencia de gen directa (Paszowski *et al.* (1984) *EMBO J.* 3:2717-2722), y aceleración balística de partícula (ver, por ejemplo, Sanford *et al.*, Patente U.S. No. 4,945,050; Tomes *et al.* (1995) "Direct ADN Transfer into Intact Plant Cells via Microprojectile Bombardment", in *Plant Cell, Tissue, and Organ Culture: Fundamental methods*, ed. Gamborg y Phillips (Springer-Verlag, Berlin); and McCabe *et al.* (1988) *Biotechnology* 6:923-926). También ver Weissinger *et al.* (1988) *Ann Rev. Genet.* 22:421-477; Sanford *et al.* (1987) *Particulate Science and Technology* 5:27-37 (onion); Christou *et al.* (1988) *Plant Physiol.* 87:671-674 (soybean); McCabe *et al.* (1988) *Bio/Technology* 6:923-926 (soybean); Finer y McMullen (1991) *En Vitro Cell Dev. Biol.* 27P:175-182 (soybean); Singh *et al.* (1998) *Theor. Appl. Genet.* 96:319-324 (soybean); Datta *et al.* (1990) *Biotechnology* 8:736-740 (arroz); Klein *et al.* (1988) *Proc. Natl. Acad. Sci. USA* 85:4305-4309 (maíz); Klein *et al.* (1988) *Biotechnology* 6:559-563 (maíz); Tomes, Patente U.S. No. 5,240,855; Buising *et al.*, Patente U.S. Nos. 5,322,783 y 5,324,646; Tomes *et al.* (1995) "Direct ADN Transfer into Intact Plant Cells via Microprojectile Bombardment", in *Plant Cell, Tissue, y Organ Culture: Fundamental métodos*, ed. Gamborg (Springer-Verlag, Berlin) (maíz); Klein *et al.* (1988) *Plant Physiol.* 91:440-444 (maíz); Fromm *et al.* (1990) *Biotechnology* 8:833-839 (maíz); Hooykaas-Van Slogteren *et al.* (1984) *Nature (London)* 311:763-764; Bytebier *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:5345-5349 (Liliaceae); De Wet *et al.* (1985) in *The Experimental Manipulation of Ovule Tissues*, ed. Chapman *et al.* (Longman, New York), pp. 197-209 (pollen); Kaeppler *et al.* (1990) *Plant Cell Reports* 9:415-418 y Kaeppler *et al.* (1992) *Theor. Appl. Genet.* 84:560-566 (whisker-mediated transformación); D'Halluin *et al.* (1992) *Plant Cell* 4:1495-1505 (electroporación); Li *et al.* (1993) *Plant Cell Reports* 12:250-255 y Christou y Ford (1995) *Annals de Botany* 75: 407-413 (arroz); Osjoda *et al.* (1996) *Nature Biotechnology* 14:745-750 (maize via *Agrobacterium tumefaciens*);

Como alternativa, las secuencias de nucleótidos de la invención se pueden introducir en un organismo y permitir experimentar la recombinación con regiones homólogas del genoma del organismo. Tales metodologías de recombinación de homólogos son bien conocidos por aquellos de ordinaria habilidad en el oficio y se pueden utilizar para incorporar establemente las secuencias de la invención en un organismo. Además, tales estrategias se pueden utilizar

para introducir “mutaciones bloqueadas” en un gen específico de un organismo que comparte sustancial homología con las secuencias de la invención. Una mutación noqueada es cualquier mutación en la secuencia de un gen que elimina o reduce considerablemente la función o el nivel del producto codificado por el gen. Los métodos que involucran la transformación de un organismo seguido por recombinación de homólogos para integrar establemente las secuencias de la invención en el organismo genoma se abarcan por la invención. La invención se dirige particularmente a los métodos donde las secuencias de la invención se utilizan para alterar el crecimiento de un organismo. Tales métodos abarcan el uso de las secuencias de la invención para interferir con la función o síntesis de una P-glicoproteína que controla el crecimiento de un organismo.

Las células que han sido transformadas se pueden cultivar en las plantas de acuerdo con formas convencionales. Ver, por ejemplo, McCormick *et al.* (1986) Plant Cell Reports 5:81-84. Estas plantas luego se pueden cultivar, y polinizar con la misma cepa transformada o cepas diferentes, y el híbrido resultante que tiene expresión constitutiva de la deseada característica fenotípica identificada. Dos o más generaciones se pueden cultivar para asegurar que la expresión constitutiva de la deseada característica fenotípica se mantiene establemente y hereda y así las cosas, se han logrado las semillas cosechadas para asegurar la expresión constitutiva de la deseada característica fenotípica.

La presente invención se pueden utilizar para la transformación de cualesquiera especies de plantas, incluyendo, pero no limitando a, maíz (*Zea mays*), *Brassica sp.* (por ejemplo, *B. napus*, *B. rapa*, *B. juncea*), particularmente aquellas especies *Brassica* útiles como fuentes de aceite de semillas, alfalfa (*Medicago sativa*), arroz (*Oryza sativa*), centeno (*Secale cereale*), sorgo (*Sorghum bicolor*, *Sorghum vulgare*), mijo (por ejemplo, mijo perla (*Pennisetum glaucum*), mijo proso (*Panicum miliaceum*), mijo foxtail (*Setaria italica*), mijo finger (*Eleusine coracana*)), girasol (*Helianthus annuus*), cártamo (*Carthamus tinctorius*), trigo (*Triticum aestivum*), soja (*Glycine max*), tabaco (*Nicotiana tabacum*), patata (*Solanum tuberosum*), maní (*Arachis hypogaea*), algodón (*Gossypium barbadense*, *Gossypium hirsutum*), patata dulce (*Ipomoea batatas*), yuca (*Manihot esculenta*), café (*Coffea spp.*), coco (*Cocos nucifera*), piña (*Ananas comosus*), árboles cítricos (*Citrus spp.*), cacao (*Theobroma cacao*), té (*Camellia sinensis*), banano (*Musa spp.*), aguacate (*Persea americana*), higo (*Ficus casica*), guayaba (*Psidium guajava*), mango (*Mangifera indica*), olivo (*Olea europaea*), papaya (*Carica papaya*), anacardo (*Anacardium occidentale*), macadamia (*Macadamia integrifolia*), almendra (*Prunus amygdalus*), remolacha azucarera (*Beta vulgaris*), caña de azúcar (*Saccharum spp.*), avena, cebada, legumbres, ornamentales, y coníferas.

Las legumbres incluyen tomates (*Lycopersicon esculentum*), lechuga (por ejemplo, *Lactuca sativa*), frijoles verdes (*Phaseolus vulgaris*), habas (*Phaseolus limensis*), guisante (*Lathyrus spp.*), y miembros del género *Cucumis* tal como pepino (*C. sativus*), cantalupo (*C. cantalupensis*), y melón (*C. melo*). Las ornamentales incluyen azalea (*Rhododendron spp.*), hortensia (*Macrophylla hydrangea*), hibisco (*Hibiscus rosasanensis*), rosas (*Rosa spp.*), tulipán (*Tulipa spp.*), narciso (*Narcissus spp.*), petunias (*Petunia hybrida*), clavel (*Dianthus caryophyllus*), poinsetia (*Euphorbia pulcherrima*), y crisantemo. Las coníferas que se pueden emplear en realizar la presente invención incluyen, por ejemplo, pinos tal como pino contorta (*Pinus taeda*), pino elliotii (*Pinus elliotii*), pino ponderosa (*Pinus ponderosa*), pino lodgepole (*Pinus contorta*), y pino Monterey (*Pinus radiata*); abeto-Douglas (*Pseudotsuga menziesii*); abeto Western (*Tsuga canadensis*); abeto rojo Sitka (*Picea glauca*); secoya (*Sequoia sempervirens*); abetos auténticos tales como abeto de plata (*Abies amabilis*) y abeto bálsamo (*Abies balsamea*); y cedros tales como cedro rojo Western (*Thuja plicata*) y cedro amarillo Alaska (*Chamaecyparis nootkatensis*). Preferiblemente, las plantas de la presente invención son plantas de cosecha (por ejemplo, arroz, maíz, alfalfa, girasol, *Brassica*, soja, algodón, cártamo, maní, sorgo, trigo, mijo, tabaco, etc.), más preferiblemente plantas de maíz, arroz y de sorgo.

Los siguientes ejemplos se ofrecen a manera de ilustración y no de limitación.

Ejemplo I

El gen de enanismo del sorgo, Dw3, codifica una P-glicoproteína homóloga

Es bien establecido que el fenotipo de enanismo del sorgo conferido por la mutación recesiva *dw3* es inestable, aunque el mecanismo responsable de su inestabilidad permanece desconocido. El alelo *dw3*, refiriéndose aquí como el alelo referencial (*dw3-ref*), revierte de nuevo a la forma tipo-salvaje (conferir un fenotipo alto) con una frecuencia de aproximadamente 0.4% a aproximadamente 1%. Como un resultado, es una trivialidad para presenciar un número de plantas altas de sorgo en un campo de enanos *dw3*. Para determinar si existe cualquier relación entre el gen *br2* del maíz y el gen *dw3* del sorgo, se colectaron muestras de hojas de 8 plantas enanas y 8 altas (revertiente); estos se anticiparon para ser isogénicos *dw3* auténticos (idénticos durante todo el genoma excepto en el locus *dw3*). El ADN de estas muestras se extrajo, digirió con PstI, y sometió a un análisis Southern blot utilizando una sonda del gen *Br2* del maíz. Un polimorfismo del ADN claro y consistente se observó entre las plantas altas y enanas, con el fragmento de restricción del alelo revertiente que es aproximadamente 1.0 kb más pequeño que el alelo *dw3-ref*.

Dos conclusiones se hicieron de este resultado. Primero, el locus *Dw3* del sorgo es estructural y funcionalmente homólogo al gen *Br2* del maíz, sugiriendo que ellos pueden reunirse para ser auténticos ortólogos (i.e., derivado del mismo gen ancestral por descendencia vertical). Segundo, dado que todas las revertientes tuvieron la misma plantilla RFLP, y que el tamaño del alelo revertiente fue más pequeño que el alelo mutante, el alelo *dw3-ref* mutable se causó probablemente por una inserción. Para dirigir la última interpretación, ADN de sorgo en la cercanía del polimorfismo del *Br2*-detectado se sometió a amplificación PCR utilizando dos cebadores del oligonucleótido (SEQ ID NOS: 5-6) derivados de la secuencia de nucleótidos del gen *Br2* del maíz.

Los productos PCR se amplificaron del ADN genómico aislado de revertientes altas y plantas enanas con el alelo *dw3-ref*. Los productos PCR posteriormente se clonaron y secuenciaron. Los resultados obtenidos mostraron que una duplicación de 882 bp ha ocurrido en el exón 5 del gen *dw3* que conduce a la generación del alelo *dw3-ref* mutante (SEQ ID NO: 1). Así, el fenotipo enano *dw3* en sorgo se debe probablemente a una mutación inserción-inducida que ocurre dentro del alelo *Dw3* para dar origen al alelo *dw3-ref*. Una secuencia parcial del alelo revertiente alto, designada *Dw3-T* se revela en la SEQ ID NO: 3. La presente duplicación en el alelo *dw3-ref* mutante también parece responsable de la inestable naturaleza de *dw3-ref*. Por un mecanismo no determinado esta duplicación se elimina en las revertientes altas de *dw3-ref*.

La comparación de la secuencia parcial de aminoácidos de la proteína codificada por *Dw3-T* (SEQ ID NO: 4) revelaron que, como BR2, esta proteína pertenece a la familia de resistencia-multifármaco-como P-glicoproteínas. Mientras que esta muestra más del 96% de aminoácido de identidad con el maíz pgp1 (el gen Br2), muestra 81% y 79% de identidad con los genes de la P-glicoproteína de *Arabidopsis thaliana* y patata respectivamente.

Dado que la inestabilidad del alelo *dw3-ref* puede resultar de alguna recombinación genética entre dos copias de la parte duplicada del gen, no podría siempre ser precisa. Algunos ejemplos pueden ocurrir donde uno o más pares de base extras pueden ser dejadas atrás o suprimidas, conduciendo en cualquier caso a una mutación de cambio del marco. Tales eventos, de esta manera se esperan que generen alelos mutantes nuevos de *dw3* que estan desprovistos de la duplicación. Y dado que la duplicación parece responsable para la inestabilidad de *dw3-ref*, los alelos nuevos mutantes de *dw3* se esperan que muestren un fenotipo de enanismo estable. Tales alelos enanos estables son altamente deseables para cosechar, cultivos de sorgo mejorados, pues la inestabilidad de *dw3* ha sido una constante nemesi para los criadores para mejorar la producción de sorgo.

Para identificar un alelo *dw3* estable, se extrajo ADN de 200 plantas de sorgo enanos y se sometio a análisis Southern blot utilizando una sonda del gen *Br2* del maíz. Se identificaron, dos plantas enanas que exhiben una plantilla de restricción que fue diferente del resto de las plantas enanas. El ADN genómico se aisló de una de estas dos plantas enanas y amplificó utilizando los cebadores del oligonucleótido (SEQ ID NOS: 5-6) según lo descrito *supra*. El producto PCR se clonó y secuenció. La comparación de la secuencia de nucleótidos del producto PCR clonado (SEQ ID NO: 2) a partir de esta planta enana para la secuencia de *dw3-ref* (SEQ ID NO: 1) revelada que la presente duplicación en *dw3-ref* se perdió. Así, esta planta enana posee un nuevo alelo *dw3*, designado como *dw3-1*. La comparación de la secuencia de nucleótidos del alelo *dw3-1* con el alelo *Dw3-T* demostró que el nuevo alelo *dw3-1* ha experimentado cambios menores.

Para separar el alelo nuevo *dw3-1* a partir del alelo *dw3-ref* parental, la planta enana que posee el *dw3-1* se auto polinizó y las semillas de la planta se colectaron y plantaron. A partir de la progenie, las plantas que fueron homocigóticas para *dw3-1* se identificaron por un análisis Southern blot, y las plantas homocigóticas se están propagando para desarrollar el germoplasma de enanismo estable para sorgo. Además, ocho hermanas del sorgo propiedad Pioneer separadas también están siendo genotipadas para la presencia de nuevos mutantes derivados del *dw3-ref*. Las hermanas que se utilizaron fueron AGK1G, MK7G, MQC100G, ZYL24, YYU28W, CAJ14W, FYL14W, y YGC87W. Se seleccionaron con base en su frecuencia de reversión, que fueron evaluadas alta, moderada o baja. Estos consanguíneos se plantaron al aire libre en Johnston, Iowa durante el verano de 1999. Doscientas plantas de cada línea fueron genotipadas RFLP digiriendo su ADN con PstI e hibridizando las manchas resultantes con una sonda de gen específico del extremo 3' del gen *br2* del maíz. Cuatro plantas homocigóticas enanas estables se identificaron a partir de YYU28W y diez de tales plantas se identificaron a partir de FYL14W. Se han cosechado, las semillas de estas plantas enanas estables. La progenie de estas plantas enanas estables se puede utilizar directamente para la producción de híbridos de sorgo de alto rendimiento con el fenotipo enano estable deseado.

Ejemplo 2

Secuencia de Nucleótidos de un Gen *Dw3* que Codifica un Producto del Gen Funcional

Con el fin de clonar la secuencia total de ambos alelos, el funcional (*Dw3*) y el mutante (*dw3-ref*) del locus *dw3*, una planta revertiente alta y una hermana enana se seleccionaron a partir del AGK1G congénito. En esta línea congénita, las plantas altas aparecen con una frecuencia de 0.1-0.4%. Se esperaba que el genotipo de la planta revertiente alta fuera heterocigota en el locus *dw3* pero por otra parte idéntica a su hermana enana por todo el genoma. Para confirmar que la revertiente alta era heterocigota en el locus *dw3*, las muestras de ADN aisladas a partir de esta planta y un número de hermanas enanas se caracterizaron por un análisis Southern utilizando tres sondas que representan las partes 5', intermedio y 3' del gen del maíz *br2*. Como se esperaba, el polimorfismo entre hermanas enanas y la planta alta se localizó solo en el extremo 3' de *dw3*. Este análisis permitió la identificación de dos fragmentos EcoRI a partir de la revertiente alta que, cuando se combinó contenía el alelo *Dw3* completo. Estos fueron, un fragmento EcoRI de 14kb que contenía la porción 5' del gen y un fragmento de 8.1 kb que contenía el resto del gen. Un fragmento de 9.0 kb a partir del alelo *dw3-ref* se determinó que correspondía al fragmento EcoRI de 8.1 kb del alelo *Dw3*. Tres bibliotecas tamaño-seleccionadas (que contienen los fragmentos 14 kb/EcoRI y 8.1 kb/EcoRI a partir de la revertiente alta y el fragmento 9.0 kb/EcoRI de la hermana enana) se construyeron en vectores de clonación Lambda de Stratagene. La biblioteca fragmento 14.0 kb/EcoRI se construyó en λ Dash II y se seleccionó con una sonda que proviene del terminal del extremo 5' del gen del maíz *br2*. Las otras dos bibliotecas se prepararon en λ ZapII y se seleccionaron con una sonda del extremo 3' del gen del maíz *br2*. Los clones positivos se aislaron y ADN λ se extrajo para cada uno de estos clones. Los genes *Dw3* y *dw3* se amplificaron PCR en cuatro fragmentos traslapados de 0.5 kb, 2.4 kb, 3.0 kb, y 1.3

ES 2 288 881 T3

kb utilizando cebadores específicos de genes y ADN λ como una plantilla. Estos fragmentos PCR se clonaron en el vector TOPO (Invitrogen). A partir del clon *dw3* enano de 9.0 kb, un fragmento SacI único de 888 bp que contiene una parte de la región duplicada se subclonó en el vector pBSK+ (Stratagene).

- 5 El ADN de al menos dos colonias de cada clon PCR se secuenció utilizando M13 hacia delante, M13 reverso, y los cebadores específicos de genes (GSPs). El fragmento SacI 888 bp del clon *dw3-ref* se secuenció utilizando cebadores de vector específico T3 y T7 solos. La información de la secuencia, ambos del extremo de los terminales 5' y 3' de los genes *Dw3* y *dw3*, se recolectó secuenciando directamente el ADN λ de los clones 14.0 kb y 8.1 kb, utilizando cebadores de gen específico. Toda la información de la secuencia se recopiló y comparó para dar a conocer la causa de enanismo en sorgo. En alineaciones en parejas entre los genes *Dw3* y *Br2* se hizo en el nivel de proteína utilizando el Programa Clustal W y en el nivel del nucleótido utilizando el Programa BLAST de NCBL.

15 Un polinucleótido de 6827 bp que contiene el gen *Dw3* longitud completa se ensambló y presentó en la SEQ ID NO: 7. Estructuralmente, el gen *Dw3* tiene cinco exones y cuatro intrones. La longitud de cinco exones, del exón 1 al exón 5, es 616 bp, 537 bp, 326 bp, 230 bp, y 2400 bp, respectivamente. El intrón 1 tiene 165 bp (nucleótidos 639-803 de SEQ ID NO: 7); intrón 2 tiene 110 bp (nucleótidos 1441-1550); intrón 3 tiene 846 bp (nucleótidos 1877-2722); y el intrón 4 tiene 1471 bp (nucleótidos 2953-4423) de longitud. Los límites intrón/exón de *Dw3* son idénticos a aquellos del gen *br2* de maíz. El predicho *Dw3*-cADN tiene 4209 bp de longitud del codón inicial al final del codón de terminación (SEQ ID NO: 8) y es de esta manera 28 bp más larga que la región análoga del *Br2*-cADN. De modo semejante, la proteína codificada predicha por *Dw3* tiene 1402 aminoácidos de largo (SEQ ID NO: 9), según se compara con los 1394 aminoácidos de la proteína predicha del gen de *Br2*. Los resultados de alineación múltiple muestran que el *Dw3* completo es 92% y 91.8% idéntico al gen del maíz *Br2* en el nivel del nucleótido y en el nivel de aminoácido, respectivamente.

25 El análisis PCR de la región polimórfica entre *dw3* y *br2* había sugerido primero que una duplicación de una parte de exón 5 resultó en el alelo de enanismo *dw3-ref* del sorgo. Para dirigir si era exclusivamente la razón para la naturaleza mutante del alelo *dw3-ref*, la secuencia del alelo *Dw3* de la revertiente alta fue comparada con la del alelo *dw3-ref*. Como se muestra previamente, la diferencia se detectó solo en el exón 5 entre estos alelos. En el alelo mutante (*dw3-ref*, SEQ ID NO: 1), un tramo de 882 bp en el exón 5 (a partir de los nucleótidos 5650-6531 de SEQ ID NO: 7) se duplica en la posición del nucleótido 6532 en la misma dirección. Esta duplicación convirtió el fragmento de restricción PstI de 1312 bp (de los nucleótidos 5463 bp a 6775 de la SEQ ID NO: 7) en el alelo *Dw3* funcional al fragmento PstI de 2194 bp encontrado en el alelo *dw3-ref*, y de esta manera fue la causa para el polimorfismo entre estos dos alelos. Dado que ningún otro cambio se encontró entre estos alelos, los resultados claramente implican esta duplicación como la única causa para crear el alelo de enanismo *dw3* del sorgo. La adición de 882 bp al cADN no tendrá duda una ramificación seria para la estructura y actividad de la proteína DW3. Estos hallazgos también muestran como el alelo *dw3-ref* se corrige espontáneamente, de vez en cuando, consiguiendo librarse de la duplicación. El mecanismo, por el cual esta corrección ocurre, sigue siendo desconocido, al igual que el mecanismo por que ocurre la duplicación en el primer lugar.

40 Ejemplo 3

Transformación del Maíz por Bombardeo y Regeneración de Partículas de Plantas Transgénicas

45 Los embriones de maíz inmaduros a partir del plantas donantes de invernadero se bombardearon con un plásmido que contiene una secuencia de nucleótidos de la P-glicoproteína de la invención ligada operablemente a un promotor que conduce la expresión en una planta y el gen marcador genético PAT (Wohlleben *et al.* (1988) Gene 70:25-37), que confiere resistencia al herbicida Bialaphos. Como alternativa, el gen marcador genético se suple en un plásmido separado. La transformación se realiza como sigue. La fórmula del medio continúa abajo.

50 *Preparación del Tejido Diana*

55 Las espigas se descascaran y la superficie se esteriliza en lejía 30% Clorox más detergente 0.5% Micro por 20 minutos, y se enjuagan dos veces con agua estéril. Los embriones inmaduros se remueven por incisión y se coloca el eje del embrión lado abajo (escutelo lado arriba), 25 embriones por placa, sobre medio 560Y por 4 horas y luego se alinean dentro de la zona diana de 2.5-cm en preparación para bombardeo.

Preparación de ADN

60 Se construye un plásmido vector que comprende la secuencia de nucleótidos de la P-glicoproteína de la invención ligada operablemente a la planta promotor de interés. Este plásmido de ADN más el plásmido ADN que contiene un marcador genético PAT se precipita en 1.1 μ m (diámetro promedio) gránulos de tungsteno utilizando un procedimiento de precipitación de CaCl₂ como sigue:

65 100 μ l preparado de partículas de tungsteno en agua

10 μ l (1 μ g) ADN en solución reguladora Tris EDTA (1 μ g de ADN total)

ES 2 288 881 T3

100 μ l 2.5 M CaCl_2

10 μ l 0.1 M espermidina

- 5 Cada reactivo se adiciona secuencialmente a la suspensión de partícula de tungsteno, mientras se mantiene en el multitubo para vórtex. La mezcla final se somete a ultrasonido brevemente y se deja incubar bajo constante agitación en vortex por 10 minutos. Después del periodo de precipitación, los tubos se centrifugaron brevemente, se elimina el líquido, se lava con 500 ml de etanol 100%, y se centrifuga por 30 segundos. Nuevamente el líquido se elimina, y se adicionan 105 μ l de etanol 100% al granulado de partícula tungsteno final. Para el bombardeo de partícula con
10 pistola, las partículas de tungsteno/ADN se someten a ultrasonido brevemente y 10 μ l se manchan en el centro de cada microportador y se deja secar aproximadamente 2 minutos antes del bombardeo.

Tratamiento de la partícula con pistola

- 15 Las placas de la muestra se bombardearon en el nivel #4 con pistola de partícula #HE34-1 o #HE34-2. Todas las muestras reciben un solo disparo a 650 PSI, con un total de diez alícuotas tomadas de cada tubo de las partículas/ADN preparadas.

Tratamiento Subsiguiente

- 20 Siguiendo el bombardeo, los embriones se mantuvieron en el medio 560Y por 2 días, luego se transfirieron a un medio de selección 560R que contiene 3 mg/litro Bialaphos, y se subcultivan cada 2 semanas. Después de aproximadamente 10 semanas de selección, los clones de callos selección-resistentes se transfirieron al medio 288J para iniciar la regeneración de la planta. Siguiendo la maduración del embrión somático (2-4 semanas), Los embriones
25 somáticos bien desarrollados se transfirieron al medio por germinación y se transfirieron al cuarto de cultivo iluminado. Aproximadamente 7-10 días después, las plántulas desarrolladas se transfirieron al medio 272V hormona-libre en tubos por 7-10 días hasta que las plántulas están bien establecidas. Las plantas luego se transfirieron para insertar en planos (equivalente a una maceta de 2.5") que contienen suelo plantado y se cultiva por 1 semana en una cámara de crecimiento, posteriormente crece un adicional de 1-2 semanas en el invernadero, luego se transfirieron a macetas
30 600 clásicas (1.6 galón) y se cultiva a madurez. Las plantas se monitorearon y puntuaron para fenotipo enano u otro fenotipo asociado con expresión de la secuencia de nucleótidos de la P-glicoproteína de la invención.

Bombardeo y Medio de Cultivo

- 35 Medio de bombardeo (560Y) comprende 4.0 g/l sales básicas N6 (SIGMA C-1416), 1.0 ml/l Mezcla de Vitaminas de Eriksson (1000X SIGMA-1511), 0.5 mg/l de tiamina HCl, 120.0 g/l de sacarosa, 1.0 mg/l de 2,4-D, y 2.88 g/l de L-prolina (llevar a volumen con D-I H_2O siguiendo con el ajuste a pH 5.8 con KOH); 2.0 g/l de Gelrite (adicionado después de llevar a volumen con D-I H_2O); y 8.5 mg/l de nitrato de plata (adicionado después de esterilizar el medio y enfriar a temperatura ambiente). Medio de selección (560R) comprende 4.0 g/l de sales básicas N6 (SIGMA C-
40 1416), 1.0 ml/l Mezcla de Vitaminas de Eriksson (1000X SIGMA-1511), 0.5 mg/l de tiamina HCl, 30.0 g/l de sacarosa, y 2.0 mg/l de 2,4-D (llevar a volumen con H_2O D-I siguiendo el ajuste a pH 5.8 con KOH); 3.0 g/l de Gelrite (adicionado después de llevar a volumen con D-I H_2O); y 0.85 mg/l de nitrato de plata y 3.0 mg/l de bialaphos (adicionando ambos después de esterilizar el medio y enfriar a temperatura ambiente).
- 45 Medio de regeneración de la planta (288J) comprende 4.3 g/l de sales MS (GIBCO 11117-074), 5.0 ml/l solución stock de vitaminas MS (0.100 g de ácido nicotínico, 0.02 g/l de tiamina HCL, 0.10 g/l de piridoxina HCL, y 0.40 g/l de glicina llevar a volumen con H_2O D-I clarificado) (Murashige and Skoog (1962) *Physiol. Plant.* 15:473), 100 mg/l de mio-inositol, 0.5 mg/l de zeatin, 60 g/l de sacarosa, y 1.0 ml/l de 0.1 mM de ácido abscísico (llevar a volumen con H_2O D-I clarificado después de ajustar a pH 5.6); 3.0 g/l de Gelrite (adicionado después de llevar a volumen con H_2O D-I);
50 y 1.0 mg/l de ácido indolacético y 3.0 mg/l de bialaphos (adicionado después de esterilizar el medio y enfriar a 60°C). El medio de hormona-libre (272V) comprende 4.3 g/l de sales MS (GIBCO 11117-074), 5.0 ml/l de solución stock de vitaminas MS (0.100 g/l de ácido nicotínico, 0.02 g/l de tiamina HCL, 0.10 g/l de piridoxina HCL, y 0.40 g/l de glicina llevar a volumen con H_2O D-I clarificado), 0.1 g/l de mio-inositol, y 40.0 g/l de sacarosa (llevar a volumen con H_2O D-I clarificado después de ajustar el pH a 5.6); y 6 g/l de bacto-agar (adicionando después de llevar a volumen
55 con H_2O D-I clarificado), se esteriliza y enfría a 60°C.

Ejemplo 4

Transformación del Agrobacterium-Mediado del Maíz y Regeneración de Plantas Transgénicas

- 60 Para la transformación del *Agrobacterium-mediado* de maíz con una secuencia de nucleótidos de la P-glicoproteína de la invención, preferiblemente se emplea el método de Zhao (Patente U.S. No. 5,981,840, y la publicación de la patente PCT WO98/32326; los contenidos de esta se incorporan aquí por referencia). Brevemente, los embriones inmaduros se aíslan del maíz y los embriones se someten a contacto con una suspensión de *Agrobacterium*, donde las
65 bacterias son capaces de transferir la secuencia de nucleótidos de la P-glicoproteína de la invención a al menos una célula de al menos uno de los embriones inmaduros (etapa 1: la etapa de infección). En esta etapa los embriones inmaduros preferiblemente se sumergen en una suspensión de *Agrobacterium* para iniciar la inoculación. Los embriones se co-cultivan por un tiempo con el *Agrobacterium* (etapa 2: la etapa de co-cultivación). Preferiblemente los embriones

inmaduros se cultivan en un medio sólido siguiendo la etapa de infección. Continuando este periodo de co-cultivación, se contempla una etapa opcional “reposo”. En esta etapa de reposo, los embriones se incuban en la presencia de al menos un antibiótico conocido para inhibir el crecimiento del *Agrobacterium* sin la adición de un agente selectivo para los transformantes de la planta (etapa 3: etapa de reposo). Preferiblemente los embriones inmaduros se cultivan sobre un medio sólido con antibiótico, pero sin un agente de selección, para la eliminación del *Agrobacterium* y para una fase de reposo para las células infectadas. Después, los embriones inoculados se cultivan en medio que contiene un agente selectivo y se recupera el callo transformado del cultivo (etapa 4: la etapa de selección). Preferiblemente, los embriones inmaduros se cultivan en medio sólido con un agente selectivo resultando en el crecimiento selectivo de células transformadas. El callo luego se regenera en las plantas (etapa 5: la etapa de regeneración), y preferiblemente los callos crecidos en el medio selectivo se cultivan en medio sólido para regenerar las plantas.

Aunque la precedente invención se ha descrito con algún detalle por medio de la ilustración y ejemplo, para propósitos de claridad de comprensión, será obvio que ciertos cambios y modificaciones se puedan practicar dentro del alcance de las reivindicaciones anexas.

Referencias citadas en la descripción

Esta lista de referencias citadas por el aspirante es únicamente para conveniencia del lector. No forma parte del documento de la Patente Europea. Aún cuando se ha tomado gran cuidado en recopilar las referencias, los errores u omisiones no se pueden excluir y la EPO niega toda la responsabilidad en esta consideración.

Documentos de la patente citados en la descripción

- US 4873192 A
- EP 75444 A
- US 5605793 A
- US 5837458 A
- US 5565350 A
- US 5731181 A
- US 5756325 A
- US 5760012 A
- US 5795972 A
- US 5871984 A
- WO 9849350 A
- WO 9907865 A
- WO 9925821 A
- US 5380831 A
- US 5436391 A
- US 5283184 A
- US 5034323 A
- WO 9943838 A
- US 6072050 A
- US 409297 A
- US 5608149 A
- US 5608144 A

- US 5604121 A
- US 5569597 A
- 5 • US 5466785 A
- US 5399680 A
- US 5268463 A
- 10 • US 5608142 A
- US 5837876 A
- 15 • US 5750386 A
- US 5633363 A
- US 5459252 A
- 20 • US 5401836 A
- US 5110732 A
- 25 • US 5023179 A
- US 07610098 P
- US 07964898 P
- 30 • WO 9943819 A
- US 5428148 A
- 35 • US 5814618 A
- US 5789156 A
- US 5563055 A, Townsend
- 40 • US 5981840 A, Zhao
- US 4945050 A, Sanford
- 45 • US 5240855 A, Tomes
- US 5322783 A, Buising
- US 5324646 A
- 50 • WO 9832326 A

Literatura no-patente citada en la descripción

- 55 • **QUAIL**, P.H. *Science*, 1995, vol. 268, 675-680
- **KENDE et al.** *Plant Cell*, 1997, vol. 9, 1197-1210
- **YOKOTA**, T. *Trends Plant Sci.*, 1997, vol. 2, 137-143
- 60 • **AZPIROZ et al.** *Plant Cell*, 1998, vol. 10, 219-230
- **CHOE et al.** *Plant Cell*, 1998, vol. 10, 231-243
- 65 • **KLEE et al.** *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, 1991, vol. 42, 529-551
- **COE et al.** *Corn & Corn Improvement.*, 1988

ES 2 288 881 T3

- **SHERIDAN**, W.F. *Annu. Rev. Genet.*, 1988, vol. 22, 353-385
- **PHINNEY et al.** *Curr. Top. Plant Biochem. Physiol.*, 1985, vol. 4, 67-74
- 5 • **FUJIOKA et al.** *Proc. Natl. Acad. Sci. USA*, 1988, vol. 85, 9031-9035
- **BENSEN et al.** *Plant Cell*, 1995, vol. 7, 75-84
- **WINKLER et al.** *Plant Cell*, 1995, vol. 7, 1307-1317
- 10 • **KEMPTON**, J.H. *J. Hered.*, 1920, vol. 11, 111-115
- **COE et al.** *Corn & Corn Improvement*, 1988, vol. 22, 353-385
- 15 • **PENDLETON et al.** *Crop Sci.*, 1961, vol. 1, 433-435
- **DUVICK**, D.N. *Maydica*, 1977, vol. 22, 187-196
- **DJISBAR et al.** *Maydica*, 1987, vol. 32, 107-123
- 20 • **RUSSEL**, W.A. *Adv. Agron.*, 1991, vol. 46, 245-298
- **KUNKEL**. *Proc. Natl. Acad. Sci. USA*, 1985, vol. 82, 488-492
- 25 • **KUNKEL et al.** *Methods in Enzymol.*, 1987, vol. 154, 367-382
- Techniques in Molecular Biology. MacMillan Publishing Company, 1983
- **DAYHOFF et al.** Atlas of Protein Sequence and Structure. *Natl. Biomed. Res. Found.*, 1978
- 30 • **STEMMER**. *Proc. Natl. Acad. Sci. USA*, 1994, vol. 91, 10747-10751
- **STEMMER**. *Nature*, 1994, vol. 370, 389-391
- 35 • **CRAMERI et al.** *Nature Biotech.*, 1997, vol. 15, 436-438
- **MOORE et al.** *J. Mol. Biol.*, 1997, vol. 272, 336-347
- **ZHANG et al.** *Proc. Natl. Acad. Sci. USA*, 1997, vol. 94, 4504-4509
- 40 • **CRAMERI et al.** *Nature*, 1998, vol. 391, 288-291
- **SAMBROOK et al.** Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, 1989
- 45 • PCR Protocols: A Guide to Methods and Applications. Academic Press, 1990
- PCR Strategies. Academic Press, 1995
- PCR Methods Manual. Academic Press, 1999.
- 50 • **MEINKOTH; WAHL**. *Anal. Biochem.*, 1984, vol. 138, 267-284
- **TIJSEN**. Laboratory Techniques in Biochemistry and Molecular Biology-Hybridization with Nucleic Acid Probes. Elsevier, 1993, vol. I
- 55 • Current Protocols in Molecular Biology. Greene Publishing and Wiley-Interscience, 1995
- **MYERS; MILLER**. *CABIOS*, 1988, vol. 4, 11-17
- 60 • **SMITH et al.** *Adv. Appl. Math.*, 1981, vol. 2, 482
- **WUNSCH**. *J. Mol. Biol.*, 1970, vol. 48, 443-453
- **PEARSON; LIPMAN**. *Proc. Natl. Acad. Sci.*, 1988, vol. 85, 2444-2448
- 65 • **KARLIN; ALTSCHUL**. *Proc. Natl. Acad. Sci. USA*, 1990, 872264
- **KARLIN; ALTSCHUL**. *Proc. Natl. Acad. Sci. USA*, 1993, vol. 90, 5873-5877

- **HIGGINS** *et al. Gene*, 1988, vol. 73, 237-244
- **HIGGINS** *et al. CABIOS*, 1989, vol. 5, 151-153
- 5 • **CORPET** *et al. Nucleic Acids Res.*, 1988, vol. 16, 10881-90
- **HUANG** *et al. CABIOS*, 1992, vol. 8, 155-65
- **PEARSON** *et al. Meth. Mol. Biol.*, 1994, vol. 24, 307-331
- 10 • **ALTSCHUL** *et al. J. Mol. Biol.*, 1990, vol. 215, 403
- **ALTSCHUL** *et al. Nucleic Acids Res.*, 1997, vol. 25, 3389
- 15 • **NEEDLEMAN; WUNSCH.** *J. Mol. Biol.*, 1970, vol. 48, 443-453
- **HENIKOFF; HENIKOFF.** *Proc. Natl. Acad. Sci. USA*, 1989, vol. 89, 10915
- **NEEDLEMAN** *et al. J. Mol. Biol.*, 1970, vol. 48, 443
- 20 • **BEETHAM** *et al. Proc. Natl. Acad. Sci. USA*, 1999, vol. 96, 8774-8778
- **GUERINEAU** *et al. Mol. Gen. Genet.*, 1991, vol. 262, 141-144
- 25 • **PROUDFOOT.** *Cell*, 1991, vol. 64, 671-674
- **SANFACON** *et al. Genes Dev.*, 1991, vol. 5, 141-149
- **MOGEN** *et al. Plant Cell*, 1990, vol. 2, 1261-1272
- 30 • **MUNROE** *et al. Gene*, 1990, vol. 91, 151-158
- **BALLAS** *et al. Nucleic Acids Res.*, 1989, vol. 17, 7891-7903
- 35 • **JOSHI** *et al. Nucleic Acid Res.*, 1987, vol. 15, 9627-9639
- **MURRAY** *et al. Nucleic Acids Res.*, 1989, vol. 17, 477-498
- **ELROY-STEIN** *et al. PNAS USA*, 1989, vol. 86, 6126-6130
- 40 • *Virology*, vol. 154, 9-20
- **MACEJAK** *et al. Nature*, 1991, vol. 353, 90-94
- 45 • **JOBLING** *et al. Nature*, 1987, vol. 325, 622-625
- **GALLIE** *et al. Molecular Biology of RNA*, 1989, 237-256
- **LOMMEL** *et al. Virology*, 1991, vol. 81, 382-385
- 50 • **CIOPPA** *et al. Plant Physiol.*, 1987, vol. 84, 965-968
- **YARRANTON.** *Curr. Opin. Biotech.*, 1992, vol. 3, 506-511
- 55 • **CHRISTOPHERSON** *et al. Proc. Natl. Acad. Sci. USA*, 1992, vol. 89, 6314-6318
- **YAO** *et al. Cell*, 1992, vol. 71, 63-72
- **REZNIKOFF.** *Mol. Microbiol.*, 1992, vol. 6, 2419-2422
- 60 • **BARKLEY** *et al. The Operon*, 1980, 177-220
- **HU** *et al. Cell*, 1987, vol. 48, 555-566
- 65 • **BROWN** *et al. Cell*, 1987, vol. 49, 603-612
- **FIGGE.** *Cell*, 1988, vol. 52, 713-722

ES 2 288 881 T3

- **DEUSCHLE** *et al. Proc. Natl. Acad. Sci. USA*, 1989, vol. 86, 5400-5404
- **FUERST** *et al. Proc. Natl. Acad. Sci. USA*, 1989, vol. 86, 2549-2553
- 5 • **DEUSCHLE** *et al. Science*, 1990, vol. 248, 480-483
- **GOSSEN**. *Ph.D. Thesis*, 1993
- **REINES** *et al. Proc. Natl. Acad. Sci. USA*, 1993, vol. 90, 1917-1921
- 10 • **LABOW** *et al. Mol. Cell. Biol.*, 1990, vol. 10, 3343-3356
- **ZAMBRETTI** *et al. Proc. Natl. Acad. Sci. USA*, 1992, vol. 89, 3952-3956
- 15 • **BAIM** *et al. Proc. Natl. Acad. Sci. USA*, 1991, vol. 88, 5072-5076
- **WYBORSKI** *et al. Nucleic Acids Res.*, 1991, vol. 19, 4647-4653
- **HILLENAND-WISSMAN**. *Topics Mol. Struc. Biol.*, 1989, vol. 10, 143-162
- 20 • **DEGENKOLB** *et al. Antimicrob. Agents Chemother.*, 1991, vol. 35, 1591-1595
- **KLEINSCHNIDT** *et al. Biochemistry*, 1988, vol. 27, 1094-1104
- 25 • **BONIN**. *Ph.D. Thesis*, 1993
- **GOSSEN** *et al. Proc. Natl. Acad. Sci. USA*, 1992, vol. 89, 5547-5551
- **OLIVA** *et al. Antimicrob. Agents Chemother.*, 1992, vol. 36, 913-919
- 30 • **HLAVKA** *et al. Handbook of Experimental Pharmacology*. Springer-Verlag, 1985, vol. 78
- **GILL** *et al. Nature*, 1988, vol. 334, 721-724
- 35 • **ODELL** *et al. Nature*, 1985, vol. 313, 810-812
- **MCELROY** *et al. Plant Cell*, 1990, vol. 2, 163-171
- **CHRISTENSEN** *et al. Plant Mol. Biol.*, 1989, vol. 12, 619-632
- 40 • **CHRISTENSEN** *et al. Plant Mol. Biol.*, 1992, vol. 18, 675-689
- **LAST** *et al. Theor. Appl. Genet.*, 1991, vol. 81, 581-588
- 45 • **VELTEN** *et al. EMBO J.*, 1984, vol. 3, 2723-2730
- **YAMAMOTO** *et al. Plant J.*, 1997, vol. 12 (2), 255-265
- **KAWAMATA** *et al. Plant Cell Physiol.*, 1997, vol. 38 (7), 792-803
- 50 • **HANSEN** *et al. Mol. Gen. Genet.*, 1997, vol. 254 (3), 337-343
- **RUSSELL** *et al. Transgenic Res.*, 1997, vol. 6 (2), 157-168
- 55 • **RINEHART** *et al. Plant Physiol.*, 1996, vol. 112 (3), 1331-1341
- **VAN CAMP** *et al. Plant Physiol.*, 1996, vol. 112 (2), 525-535
- **CANEVASCINI** *et al. Plant Physiol.*, 1996, vol. 112 (2), 513-524
- 60 • **YAMAMOTO** *et al. Plant Cell Physiol.*, 1994, vol. 35 (5), 773-778
- **LAM**. *Results Probl. Cell Differ.*, 1994, vol. 20, 181-196
- 65 • **OROZCO** *et al. Plant Mol Biol.*, 1993, vol. 23 (6), 1129-1138
- **MATSUOKA** *et al. Proc Natl. Acad. Sci. USA*, 1993, vol. 90 (20), 9586-9590

- GUEVARA-GARCIA *et al. Plant J.*, 1993, vol. 4 (3), 495-505
- HANSEN *et al. Mol. Gen. Genet.*, 1997, vol. 254 (3), 337-343
- 5 • RUSSELL *et al. Transgenic Res.*, 1997, (2), 157-168
- OROZCO *et al. Plant Mol. Biol.*, 1993, vol. 23 ((6)), 1129-1138
- MATSUOKA *et al. Proc. Natl. Acad. Sci. USA*, 1993, vol. 90 (20), 9586-9590
- 10 • HIRE *et al. Plant Mol. Biol.*, 1992, vol. 20 ((2)), 207-218
- KELLER; BAUMGARTNER. *Plant Cell*, 1991, vol. 3 (10), 1051-1061
- 15 • SANGER *et al. Plant Mol. Biol.*, 1990, vol. 14 (3), 433-443
- MIAO *et al. Plant Cell*, 1991, vol. 3 (1), 11-22
- BOGUSZ *et al. Plant Cell*, 1990, vol. 2 (7), 633-641
- 20 • LIMERICK. *Plant Science*, vol. 79 (1), 69-76
- EMBO J., vol. 8 (2), 343-350
- 25 • KUSTER *et al. Plant Mol. Biol.*, 1995, vol. 29 (4), 759-772
- CAPANA *et al. Plant Mol. Biol.*, 1994, vol. 25 (4), 681-691
- REDOLFI *et al. Neth. J. Plant Pathol.*, 1983, vol. 89, 245-254
- 30 • UKNES *et al. Plant Cell*, 1992, vol. 4, 645-656
- VAN LOON. *Plant Mol. Virol.*, 1985, vol. 4, 111-116
- 35 • MARINEAU *et al. Plant Mol. Biol.*, 1987, vol. 9, 335-342
- MATTON *et al. Molecular Plant-Microbe Interactions*, 1989, vol. 2, 325-331
- SOMSISCH *et al. Proc. Natl. Acad. Sci. USA*, 1986, vol. 83, 2427-2430
- 40 • SOMSISCH *et al. Mol. Gen. Genet.*, 1988, vol. 2, 93-98
- YANG. *Proc. Natl. Acad. Sci. USA*, 1996, vol. 93, 14972-14977
- 45 • CHEN *et al. Plant J.*, 1996, vol. 10, 955-966
- ZHANG *et al. Proc. Natl. Acad. Sci. USA*, 1994, vol. 91, 2507-2511
- WARNER *et al. Plant J.*, 1993, vol. 3, 191-201
- 50 • SIEBERTZ *et al. Plant Cell*, 1989, vol. 1, 961-968
- CORDERO *et al. Physiol. Mol. Plant Path.*, 1992, vol. 41, 189-200
- 55 • RYAN. *Ann. Rev. Phytopath.*, 1990, vol. 28, 425-449
- DUAN *et al. Nature Biotechnology*, 1996, vol. 14, 494-498
- STANFORD *et al. Mol. Gen. Genet.*, 1989, vol. 215, 200-208
- 60 • MCGURL *et al. Science*, 1992, vol. 225, 1570-1573
- ROHMEIER *et al. Plant Mol. Biol.*, 1993, vol. 22, 783-792
- 65 • ECKELKAMP *et al. FEBS Letters*, 1993, vol. 323, 73-76
- CORDEROK *et al. Plant J.*, 1994, vol. 6 (2), 141-150

- **SCHENA** *et al. Proc. Natl. Acad. Sci. USA*, 1991, vol. 88, 10421-10425
- **MCNELLIS** *et al. Plant J*, 1998, vol. 14 (2), 247-257
- 5 • **GATZ** *et al. Mol. Gen. Genet.*, 1991, vol. 227, 229-237
- **CROSSWAY** *et al. Biotechniques*, 1986, vol. 4, 320-334
- **RIGGS** *et al. Proc. Natl. Acad. Sci. USA*, 1986, vol. 83, 5602-5606
- 10 • **PASZKOWSKI** *et al. EMBO J.*, 1984, vol. 3, 2717-2722
- Direct DNA Transfer into Intact Plant Cells via Microprojectile Bombardment. **TOMES** *et al. Plant Cell, Tissue, and Organ Culture: Fundamental Methods*. Springer-Verlag, 1995
- 15 • **MCCABE** *et al. Biotechnology*, 1988, vol. 6, 923-926
- **WEISSINGER** *et al. Ann Rev. Genet.*, 1988, vol. 22, 421-477
- 20 • **SANFORD** *et al. Particulate Science and Technology*, 1987, vol. 5, 27-37
- **CHRISTOU** *et al. Plant Physiol.*, 1988, vol. 87, 671-674
- **MCCABE** *et al. Bio/Technology*, 1988, vol. 6, 923-926
- 25 • **FINER; MCMULLEN**. *In Vitro Cell Dev. Biol.*, 1991, vol. 27P, 175-182
- **SINGH** *et al. Theor. Appl. Genet.*, 1998, vol. 96, 319-324
- 30 • **DATTA** *et al. Biotechnology*, 1990, vol. 8, 736-740
- **KLEIN** *et al. Proc. Natl. Acad. Sci. USA*, 1988, vol. 85, 4305-4309
- **KLEIN** *et al. Biotechnology*, 1988, vol. 6, 559-563
- 35 • **KLEIN** *et al. Plant Physiol.*, 1988, vol. 91, 440-444
- **FROMM** *et al. Biotechnology*, 1990, vol. 8, 833-839
- 40 • **HOOYKAAS-VAN SLOGTEREN** *et al. Nature*, 1984, vol. 311, 763-764
- **BYTEBIER** *et al. Proc. Natl. Acad. Sci. USA*, 1987, vol. 84, 5345-5349
- **DE WET** *et al. The Experimental Manipulation of Ovule Tissues*, 1985, 197-209
- 45 • **KAEPPLER** *et al. Plant Cell Reports*, 1990, vol. 9, 415-418
- **KAEPPLER** *et al. Theor. Appl. Genet.*, 1992, vol. 84, 560-566
- 50 • **D'HALLUIN** *et al. Plant Cell*, 1992, vol. 4, 1495-1505
- **LI** *et al. Plant Cell Reports*, 1993, vol. 12, 250-255
- **CHRISTOU; FORD**. *Annals of Botany*, 1995, vol. 75, 407-413
- 55 • **OSJODA** *et al. Nature Biotechnology*, 1996, vol. 14, 745-750
- **MCCORMICK** *et al. Plant Cell Reports*, 1986, vol. 5, 81-84
- 60 • **WOHLLEBEN** *et al. Gene*, 1988, vol. 70, 25-37
- **MURASHIGE; SKOOG**. *Physiol. Plant.*, 1962, vol. 15, 473

REIVINDICACIONES

1. Una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos seleccionada del grupo que consiste de:

- (a) una secuencia de nucleótidos publicada en la SEQ ID NO: 1;
- (b) una secuencia de nucleótidos publicada en la SEQ ID NO: 2;
- (c) una secuencia de nucleótidos publicada en la SEQ ID NO: 3;
- (d) una secuencia de nucleótidos publicada en la SEQ ID NO: 7;
- (e) una secuencia de nucleótidos publicada en la SEQ ID NO: 8;
- (f) una secuencia de nucleótidos que consiste de al menos 150 nucleótidos contiguos de la secuencia de nucleótidos publicada en cualquiera de (a)-(e), en donde dicha secuencia de nucleótidos codifica una P-glicoproteína que controla el crecimiento de la planta;
- (g) una secuencia de nucleótidos que codifica la secuencia de aminoácidos publicada en la SEQ ID NO: 9;
- (h) una secuencia de nucleótidos que codifica al menos 70 aminoácidos contiguos de la secuencia de aminoácidos publicada en la SEQ ID NO: 9, en donde dicha secuencia de nucleótidos codifica una P-glicoproteína que controla el crecimiento de la planta;
- (i) una secuencia de nucleótidos que comprende al menos 80% de identidad a la secuencia publicada en la SEQ ID NO: 7, en donde dicha secuencia de nucleótidos codifica una P-glicoproteína que controla el crecimiento de la planta;
- (j) una secuencia de nucleótidos que comprende al menos 80% de identidad a la secuencia publicada en la SEQ ID NO: 8, en donde dicha secuencia de nucleótidos codifica una P-glicoproteína que controla el crecimiento de la planta;
- (k) una secuencia de nucleótidos que es complementaria a la secuencia de nucleótidos de cualquiera de (a)-(k); y
- (l) una secuencia de nucleótidos que hibridiza bajo condiciones rigurosas a cualquiera de la secuencia de nucleótidos de (a)-(e), o a una secuencia complementaria de estas, en donde dichas condiciones rigurosas comprenden la hibridación en una solución que comprende 50% de formamida, NaCl 1 M, 1% de SDS a 37°C y un lavado en 0.1X SSC a 60°C, y dicha secuencia de nucleótidos codifica una P-glicoproteína que controla el crecimiento de la planta.

2. Un casete de expresión que comprende la secuencia de nucleótidos de la reivindicación 1, dicha secuencia de nucleótidos ligada operablemente a un promotor que conduce la expresión en una célula de la planta.

3. El casete de expresión de la reivindicación 2, en donde dicho promotor se selecciona del grupo que consiste de los promotores de tejido-preferido, constitutivos, regulables químicamente, y de patógeno-preferido.

4. Una planta que ha sido transformado establemente con el casete de expresión de la reivindicación 2.

5. La planta de la reivindicación 4, en donde dicho promotor se selecciona del grupo que consiste de los promotores de tejido-preferido, constitutivos, regulables químicamente, y patógeno-preferido.

6. La planta de la reivindicación 4, en donde dicha secuencia de nucleótidos se liga operablemente a dicho promotor para la producción de transcripciones antisentido.

7. La planta de la reivindicación 4, en donde dicha planta es una monocotiledónea.

8. La planta de la reivindicación 7, en donde dicha monocotiledónea se selecciona del grupo que consiste de maíz, trigo, arroz, arroz Basmati, sorgo, centeno, mijo y cebada.

9. La planta de la reivindicación 4, en donde dicha planta es una dicotiledónea.

10. La planta de la reivindicación 9, en donde dicha dicotiledónea se selecciona del grupo que consiste de soja, girasol, cártamo, alfalfa, Brassica sp., algodón, maní y árboles frutales.

ES 2 288 881 T3

11. Una semilla de cualquiera de las plantas de las reivindicaciones 4-10, en donde dicha semilla comprende dicho casete de expresión.

12. Un método para modificar el crecimiento de una planta, dicho método que comprende la transformación de una planta con una secuencia de nucleótidos que codifica una P-glicoproteína en donde la dicha P-glicoproteína funciona para controlar el crecimiento de un organismo, dicha secuencia de nucleótidos ligada operablemente para la producción de las transcripciones antisentido a un promotor capaz de conducir la expresión de dicha secuencia en dicha planta; en donde dicha secuencia de nucleótidos es como se define en la reivindicación 1.

13. El método de la reivindicación 12, en donde se reduce la altura de dicha planta.

14. El método de la reivindicación 12, en donde dicha planta es una monocotiledónea.

15. El método de la reivindicación 12, en donde dicha monocotiledónea se selecciona del grupo que consiste de maíz, trigo, arroz, arroz Basmati, sorgo, centeno, mijo y cebada.

16. Una célula de la planta que ha sido transformada establemente con el casete de expresión de la reivindicación 2.

17. Un método para la identificación de una planta de sorgo que tiene un fenotipo enano estable que comprende:

(a) obtención de una planta de sorgo enano que es una descendiente de una planta de sorgo enano *dw3* inestable.

(b) análisis del ADN genómico a partir de dicha planta de sorgo enano para determinar si al menos una copia de un gen *dw3* carece de una inserción, o al menos de una porción de dicha inserción;

(c) selección de dicha planta de sorgo enano si dicha planta de sorgo enano contiene en su genoma al menos una copia de dicho gen *dw3* carente de dicha inserción, o carente de dicha porción de esta.

18. El método de la reivindicación 17, en donde dicha inserción es una duplicación o un transposón.

19. Una proteína aislada que comprende un miembro seleccionado del grupo que consiste de:

(a) un polipéptido que comprende la secuencia de aminoácidos publicada en la SEQ ID NO: 9;

(b) un polipéptido que comprende al menos 70 aminoácidos contiguos de la secuencia de aminoácidos publicada en la SEQ ID NO: 9, en donde dicho polipéptido es una P-glicoproteína que controla el crecimiento de la planta;

(c) un polipéptido codificado por la secuencia de nucleótidos publicada en la SEQ ID NO: 8; y

(d) un polipéptido codificado por una secuencia de aminoácidos que comprende al menos 80% de identidad a la secuencia amino de (a), en donde dicho polipéptido es una P-glicoproteína que controla el crecimiento de la planta.

ES 2 288 881 T3

LISTA DE SECUENCIAS

<110> Johal, Gurmukh S
Multani, Dilbag S

<120> GENES DE ENANISMO DEL SORGO Y MÉTODOS DE USO

<130> 5718-100 (035718/205458)

<140>

<141>

<150> 60/165,176

<151> 1999-11-12

<160> 9

<170> PatentIn Ver. 2.1

<210> 1

<211> 2139

<212> ADN

<213> Sorgo bicolor

<400> 1

```

ctcctcgccg tgttcccgt cgtcgtgggc gccaccgtgc tgcagaagat gttcatgaag 60
ggcttctcgg gggacctgga ggcgcgcgac gccagggcca cgcagatcgc gggcgaggcc 120
gtggccaacc tgcgcaccgt ggcgcggttc aacgcggagc gcaagatcac ggggctgttc 180
gaggccaacc tgcgcggccc gctccggcgc tgcctctgga aggggcagat cgcgcggcagc 240
ggctacggcg tggcgaggtt cctgctgtac gcgtectacg cgtcggggtt gtggtacgcg 300
gcgtggctgg tgaagcacgg cgtgtccgac ttctcgcgca ccatccgcgt gttcatggtg 360
ctgatgggtg ccgccaacgg cgcgcgcgag acgctgacgc tggcgccgga ctttgtcaag 420
ggcgggcgcg cgtgctggtc cgtgttcgag accatcgacc ggaaaacgga ggtggagccc 480
gacgacgtgg acgcggcgcc ggtgcccggag cgcgccaagg gcgaggtgga gctgaagcac 540
gtggacttct cgtaccctgc gcggccggac atccaggtgt tccgcgacct gaggctccgg 600
gcgcgcgcgg ggaagacgct ggcgctggtg ggtccgagcg ggtgcggcaa gagctcggtg 660
ctggcgctgg cgcagcggtt ctacgagccc auctccgggc gcgtgctcct ggacggcaag 720
acgtgcgca agtacaacct gcggcgctg cgcgcgctgg tggcggtggt gccgcaggag 780
ccgttcctgt tcgcggcgag catccacgac aacatcgctg acgggcgcga gggcgcgacg 840
gaggcgaggg tgggtggagg gcgcgactca gcgaacgcgc accggttcat ctccggcgctg 900
ccggaggggt acgggacgca agtggcgag cgcggggtgc agctgtcggg cgggcagcgg 960
cagcggtatc cgtcgcgcgc cgcgctggtg aagcaagcgg ccatcatgcc gctggacgag 1020
gcgaccagcg cgttgacccc gagtcggagc ggtggctctt cgaggccaac cttcgcggcc 1080
cgctccggcg cttgttctgg aaggggcaga tcgcgggaac ggtacgggcg tggcgagtt 1140
cttgctgacg cgtcctacgc gcttggggtt ttggtacccc gcgtggctag tgaagcacgg 1200
gtctccgact ttccgcgacc atccgggtgt tcatggtgct gatggtgtcc gccaacggcg 1260
ccgcgagac gctgacgtg gcgcgggact ttgtcaaggg cggcgcgcg atgcggtccg 1320
tgttcgagac catcgaccgg aaaacggagg tggagcccga cgacgtggac gcggcgccgg 1380
tgccggagcg gcccaggggc gaggtggagc tgaagcacgt ggacttctcg taccctcgc 1440
ggccggacat ccaggtgttc cgcgacctga gcctccgggc gcgcgcggg aagacgctg 1500
cgctggtggg tccgagcggg tgcggcaaga gctcgtgct gcgctggtg cagcggttct 1560
acgagcccac gtccggcgcg gtgctcctgg acggcaaggc cgtgcgcaag tacaacctgc 1620
ggcgctgctg gcgcgtggtg gcggtggtgc cgcaggagcc gttcctgttc gcggcgagca 1680
tccacgacaa catcgctac ggcgcgagc gcgcgacgga ggcggagggt gtggaggcgg 1740
cgacgcaggc gaacgcgcac cgttcctatc cggcgctgcc ggagggttac gggacgcagg 1800
tggcgagcg cggggtgcag ctgtcgggcg ggcaagcgca gcggtatcgc atcgcgcgcg 1860
ctggtgaagc gcggccatcg tgcgtctgga cgaggcgacc agcgcgctgg acgcccagtc 1920
ggagcggtgc gtgcaggagg cgtggagcgc cgcggggtcc gggcgacca ccatcgtggt 1980
ggcgcacccg ctggccacgg tgcgcggcgc gcacaccatc gcggtcatcg acgacggcaa 2040
ggtggcgagg caggggtcgc actcgacact gctcaagcac catcccagc ggtgctacgc 2100
gcggtatgct cagctgcagc ggtgacggg cgcggcggc

```

<210> 2

<211> 1267

<212> ADN

<213> Sorgo bicolor

ES 2 288 881 T3

<400> 2

```

5      ctccctcgccg tgttcccgt cgtcgtgggc gccaccgtgc tgcagaacat gttcatgaag 60
      ggcttctcgg gggacctgg ggcgcgcac gccagggcc cgcagatcgc gggcgaggcc 120
      gtggccaacc tgcgcaccgt ggcgcggttc aacgcggagc gcaagatcac ggggctgttc 180
      gaggccaacc tgcgcggccc gctcgggcgc tgcctctgga aggggcagat cgccggcagc 240
      ggctacggcg tggcgagtt cctgctgtac gcgtcctac cgctggggct gtggtacgcg 300
      gcgtggctgg tgaagcacgg cgtgtccgac ttctcgcgca ccatccgcgt gttcatggtg 360
10     ctgatggtgt ccgcgaacgg gcgcccgcg agacgctgac gctggcgccg gacttcatca 420
      agggcgggcg cgcgatgcgg tcggtgttcg agacgatcga ccgcaagacg gaggtggagc 480
      ccgacgacgt ggacgcggcg ccggtgccgg agcggccgag gggcgagggtg gagctgaagc 540
      acgtggactt ctctgaccgg tcgcggccgg acatccaggt gttccgcgac ctgagcctcc 600
15     gtgcgcgcgc cgggaagacg ctggcgctgg tggggccgag cgggtgcggc aagagctcgg 660
      tgctggtctt ggtgcagcgg ttctacaagc ccacgtccgg gcgctgttc ttgacggcaa 720
      agacgtgcgc aaaaacaacc ttcgggcggt ccggcgcat gttgcggtgg tacccaagaa 780
      cccgtttcct gttcgcggcg aagaatccac gagaacatcg cgcacgggcg agaggcgct 840
      acggaggcgg aggtggtgga ggcggcgggc caggcgaacg cgcaccggtt catcgcgcg 900
20     ctgcgggaag gggtagtgga cgcagaaggc cagcgcggg gtgcacctgt cggggggcag 960
      cggcagcgga tcgcgatcgc gcgcgcgctg gtgaagcagc ggccatcgtg ctgctggagc 1020
      aggcgaccag cgcgctggac gccgagtcgg agcgggtcgt gcaggaggcg ctggagcgcg 1080
      cggggtccgg gcgcaccacc atcgtggtgg cgcaccggct ggccacgggtg cgcggcgcg 1140
      acaccatcgc ggtcatcgac gacggcaagg tggcgagca ggggtcgcac tcgcacctgc 1200
25     tcaagcacca tcccgacggg tgctacgcgc ggatgctgca gctgcagcgg ctgacggggc 1260
      cggcgcg 1267

```

30 <210> 3

<211> 1261

<212> ADN

<213> Sorgo bicolor

35 <220>

<221> CDS

<222> (1)..(1245)

40 <400> 3

```

45     ctc ctc gcc gtg ttc ccg ctc gtc gtg ggc gcc acc gtg ctg cag aag 48
      Leu Leu Ala Val Phe Pro Leu Val Val Gly Ala Thr Val Leu Gln Lys
          1             5             10             15

      atg ttc atg aag ggc ttc tcg ggg gac ctg gag gcc gcg cac gcc agg 96
      Met Phe Met Lys Gly Phe Ser Gly Asp Leu Glu Ala Ala His Ala Arg
          20             25             30

      gcc acg cag atc gcg ggc gag gcc gtg gcc aac ctg cgc acc gtg gcc 144
      Ala Thr Gln Ile Ala Gly Glu Ala Val Ala Asn Leu Arg Thr Val Ala
          35             40             45

      gcg ttc aac gcg gag cgc aag atc acg ggg ctg ttc gag gcc aac ctg 192
      Ala Phe Asn Ala Glu Arg Lys Ile Thr Gly Leu Phe Glu Ala Asn Leu
          50             55             60

      cgc ggc ccg ctc cgg cgc tgc ttc tgg aag ggc cag atc gcc ggc agc 240
      Arg Gly Pro Leu Arg Arg Cys Phe Trp Lys Gly Gln Ile Ala Gly Ser
          65             70             75             80

```

65

ES 2 288 881 T3

5	ggc tac ggc gtg gcg cag ttc ctg ctg tac gcg tcc tac gcg ctg ggg Gly Tyr Gly Val Ala Gln Phe Leu Leu Tyr Ala Ser Tyr Ala Leu Gly	288
	85 90 95	
10	ctg tgg tac gcg gcg tgg ctg gtg aag cac ggc gtg tcc gac ttc tcg Leu Trp Tyr Ala Ala Trp Leu Val Lys His Gly Val Ser Asp Phe Ser	336
	100 105 110	
15	cgc acc atc cgc gtg ttc atg gtg ctg atg gtg tcc gcc aac ggc gcc Arg Thr Ile Arg Val Phe Met Val Leu Met Val Ser Ala Asn Gly Ala	384
	115 120 125	
20	gcc gag acg ctg acg ctg gcg ccg gac ttt gtc aag ggc ggg cgc gcg Ala Glu Thr Leu Thr Leu Ala Pro Asp Phe Val Lys Gly Gly Arg Ala	432
	130 135 140	
25	atg cgg tcc gtg ttc gag acc atc gac cgg aaa acg gag gtg gag ccc Met Arg Ser Val Phe Glu Thr Ile Asp Arg Lys Thr Glu Val Glu Pro	480
	145 150 155 160	
30	gac gac gtg gac gcg gcg ccg gtg ccg gag cgg ccc aag ggc gag gtg Asp Asp Val Asp Ala Ala Pro Val Pro Glu Arg Pro Lys Gly Glu Val	528
	165 170 175	
35	gag ctg aag cac gtg gac ttc tcg tac ccg tcg cgg ccg gac atc cag Glu Leu Lys His Val Asp Phe Ser Tyr Pro Ser Arg Pro Asp Ile Gln	576
	180 185 190	
40	gtg ttc cgc gac ctg agc ctc cgg gcg cgc gcc ggg aag acg ctg gcg Val Phe Arg Asp Leu Ser Leu Arg Ala Arg Ala Gly Lys Thr Leu Ala	624
	195 200 205	
45	ctg gtg ggt ccg agc ggg tgc ggc aag agc tcg gtg ctg gcg ctg gtg Leu Val Gly Pro Ser Gly Cys Gly Lys Ser Ser Val Leu Ala Leu Val	672
	210 215 220	
50	cag cgg ttc tac gag ccc acg tcc ggg cgc gtg ctc ctg gac ggc aag Gln Arg Phe Tyr Glu Pro Thr Ser Gly Arg Val Leu Leu Asp Gly Lys	720
	225 230 235 240	
55	gac gtg cgc aag tac aac ctg cgg gcg ctg cgg cgc gtg gtg gcg gtg Asp Val Arg Lys Tyr Asn Leu Arg Ala Leu Arg Arg Val Val Ala Val	768
	245 250 255	
60	gcg ccg cag gag ccg ttc ctg ttc gcg gcg agc atc cac gac aac atc Ala Pro Gln Glu Pro Phe Leu Phe Ala Ala Ser Ile His Asp Asn Ile	816
	260 265 270	
65	gcg tac ggg cgc gag ggc gcg acg gag gcg gag gtg gtg gag gcg gcg Ala Tyr Gly Arg Glu Gly Ala Thr Glu Ala Glu Val Val Glu Ala Ala	864
	275 280 285	
70	acg cag gcg aac gcg cac cgg ttc atc gcg gcg ctg ccg gag ggc tac Thr Gln Ala Asn Ala His Arg Phe Ile Ala Ala Leu Pro Glu Gly Tyr	912
	290 295 300	
75	ggg acg cag gtg ggc gag cgc ggg gtg cag ctg tcg ggc ggg cag cgg Gly Thr Gln Val Gly Glu Arg Gly Val Gln Leu Ser Gly Gly Gln Arg	960
	305 310 315 320	

ES 2 288 881 T3

	cag	cgg	atc	gcg	atc	gcg	cgc	gct	ggg	aag	cag	cgg	cca	tcg	tgc	tgc	1008
	Gln	Arg	Ile	Ala	Ile	Ala	Arg	Ala	Gly	Lys	Gln	Arg	Pro	Ser	Cys	Cys	
					325					330					335		
5	tgg	acg	agg	cga	cca	gcg	cgc	tgg	acg	ccg	agt	cgg	agc	ggg	gcg	tgc	1056
	Trp	Thr	Arg	Arg	Pro	Ala	Arg	Trp	Thr	Pro	Ser	Arg	Ser	Gly	Ala	Cys	
				340					345					350			
10	agg	agg	cgc	tgg	agc	gcg	cgg	ggg	ccg	ggc	gca	cca	cca	tcg	tgg	tgg	1104
	Arg	Arg	Arg	Trp	Ser	Ala	Arg	Gly	Pro	Gly	Ala	Pro	Pro	Ser	Trp	Trp	
			355					360					365				
15	cgc	acc	ggc	tgg	cca	cgg	tgc	gcg	gcg	cgc	aca	cca	tcg	cgg	tca	tcg	1152
	Arg	Thr	Gly	Trp	Pro	Arg	Cys	Ala	Ala	Arg	Thr	Pro	Ser	Arg	Ser	Ser	
		370					375					380					
20	acg	acg	gca	agg	tgg	cgg	agc	agg	ggg	cgc	act	cgc	acc	tgc	tca	agc	1200
	Thr	Thr	Ala	Arg	Trp	Arg	Ser	Arg	Gly	Arg	Thr	Arg	Thr	Cys	Ser	Ser	
	385					390					395					400	
25	acc	atc	ccg	acg	ggg	gct	acg	cgc	gga	tgc	tgc	agc	tgc	agc	ggc		1245
	Thr	Ile	Pro	Thr	Gly	Ala	Thr	Arg	Gly	Cys	Cys	Ser	Cys	Ser	Gly		
					405					410					415		
	tgacgggcgc	ggcggc															1261
30	<210>	4															
	<211>	415															
	<212>	PRT															
	<213>	Sorgo bicolor															
35	<400>	4															
40	Leu	Leu	Ala	Val	Phe	Pro	Leu	Val	Val	Gly	Ala	Thr	Val	Leu	Gln	Lys	
	1				5					10					15		
45	Met	Phe	Met	Lys	Gly	Phe	Ser	Gly	Asp	Leu	Glu	Ala	Ala	His	Ala	Arg	
				20					25					30			
50	Ala	Thr	Gln	Ile	Ala	Gly	Glu	Ala	Val	Ala	Asn	Leu	Arg	Thr	Val	Ala	
			35					40					45				
55	Ala	Phe	Asn	Ala	Glu	Arg	Lys	Ile	Thr	Gly	Leu	Phe	Glu	Ala	Asn	Leu	
		50					55					60					
60	Arg	Gly	Pro	Leu	Arg	Arg	Cys	Phe	Trp	Lys	Gly	Gln	Ile	Ala	Gly	Ser	
	65					70					75					80	
65	Gly	Tyr	Gly	Val	Ala	Gln	Phe	Leu	Leu	Tyr	Ala	Ser	Tyr	Ala	Leu	Gly	
					85					90					95		
70	Leu	Trp	Tyr	Ala	Ala	Trp	Leu	Val	Lys	His	Gly	Val	Ser	Asp	Phe	Ser	
				100					105					110			
75	Arg	Thr	Ile	Arg	Val	Phe	Met	Val	Leu	Met	Val	Ser	Ala	Asn	Gly	Ala	
			115</														

ES 2 288 881 T3

	145		150		155		160									
	Asp	Asp	Val	Asp	Ala	Ala	Pro	Val	Pro	Glu	Arg	Pro	Lys	Gly	Glu	Val
				165						170					175	
5	Glu	Leu	Lys	His	Val	Asp	Phe	Ser	Tyr	Pro	Ser	Arg	Pro	Asp	Ile	Gln
			180					185					190			
	Val	Phe	Arg	Asp	Leu	Ser	Leu	Arg	Ala	Arg	Ala	Gly	Lys	Thr	Leu	Ala
10			195					200					205			
	Leu	Val	Gly	Pro	Ser	Gly	Cys	Gly	Lys	Ser	Ser	Val	Leu	Ala	Leu	Val
		210					215					220				
	Gln	Arg	Phe	Tyr	Glu	Pro	Thr	Ser	Gly	Arg	Val	Leu	Leu	Asp	Gly	Lys
15		225				230				235					240	
	Asp	Val	Arg	Lys	Tyr	Asn	Leu	Arg	Ala	Leu	Arg	Arg	Val	Val	Ala	Val
				245					250						255	
	Ala	Pro	Gln	Glu	Pro	Phe	Leu	Phe	Ala	Ala	Ser	Ile	His	Asp	Asn	Ile
20				260					265					270		
	Ala	Tyr	Gly	Arg	Glu	Gly	Ala	Thr	Glu	Ala	Glu	Val	Val	Glu	Ala	Ala
			275				280						285			
	Thr	Gln	Ala	Asn	Ala	His	Arg	Phe	Ile	Ala	Ala	Leu	Pro	Glu	Gly	Tyr
25			290				295					300				
	Gly	Thr	Gln	Val	Gly	Glu	Arg	Gly	Val	Gln	Leu	Ser	Gly	Gly	Gln	Arg
		305				310				315					320	
	Gln	Arg	Ile	Ala	Ile	Ala	Arg	Ala	Gly	Lys	Gln	Arg	Pro	Ser	Cys	Cys
				325					330						335	
	Trp	Thr	Arg	Arg	Pro	Ala	Arg	Trp	Thr	Pro	Ser	Arg	Ser	Gly	Ala	Cys
35				340					345					350		
	Arg	Arg	Arg	Trp	Ser	Ala	Arg	Gly	Pro	Gly	Ala	Pro	Pro	Ser	Trp	Trp
			355					360					365			
	Arg	Thr	Gly	Trp	Pro	Arg	Cys	Ala	Ala	Arg	Thr	Pro	Ser	Arg	Ser	Ser
40			370				375					380				
	Thr	Thr	Ala	Arg	Trp	Arg	Ser	Arg	Gly	Arg	Thr	Arg	Thr	Cys	Ser	Ser
		385				390				395					400	
	Thr	Ile	Pro	Thr	Gly	Ala	Thr	Arg	Gly	Cys	Cys	Ser	Cys	Ser	Gly	
45				405					410					415		

<210> 5

<211> 26

<212> ADN

<213> Secuencia Artificial

<220>

<223> Descripción de la Secuencia Artificial: cebador del oligonucleótido designado de la secuencia del gen *Zea mays* Br2

<400> 5

ctctctgcccgtgttcccgcgtcgctgt

26

<210> 6

<211> 17

<212> ADN

<213> Secuencia Artificial

ES 2 288 881 T3

<220>

<223> Descripción de la Secuencia Artificial: cebador del oligonucleótido designado de la secuencia del gen *Zea mays* Br2

5 <400> 6

gccgcccgcg ccgtcag

17

10 <210> 7

<211> 6827

<212> ADN

15 <213> Sorgo bicolor

<400> 7

```

20      gtcctctccc ggccccgat cgtgtgtctac caacgaccgc gacgagatca gggcgcgcggt 60
      cgtcgtctct ggccccctc atgccgacga cgacgcccgc gacgagtggg cccgccccga 120
      gctcggggcc ttccacctcc cctctcccgc ccaccagcct cctggtctcc acctagccgc 180
      tgggcaccaa ccggaagctg cagcagagca acccaccacg ctccctgctg cccgcccgcac 240
      cagcgacaca tccactgctg ctggtgctgc tctctcttct ccttcgcccgc ctccgcccgc 300
25      ggctcctttg gagatggacc agccgccccaa tgccaagccg gcctcctcct ccgcccgcgc 360
      cgccggcgcc aatgacaaca agaagcccac cccgcccgcg gcgtgcgcg acctcttcgc 420
      cttcgccgac ggctcgcact gcgcgctcat gctcgtcggc acgctcggcg cgtcgtcca 480
      cggtcgtcgc ctcccgtct tctccgctt cttcgccgac ctcgctgact ccttcggctc 540
      ccacgccaac gaccgggaca ccatgggtccg cctcgtcgtc aagtacgcct tctacttct 600
      cgtcgtcggg gccgcaatct gggcgtctc atgggcaggt aaccaacgtt attcctctc 660
30      ctctctccc cctcctccc ggcactgctg ctcgctcgc gaattgtctg tcgatttggg 720
      ttggtaggcg aatcacatca gtcgtcfaat cttcatggcc catggctagc aatgagatcg 780
      acctcgaat cctcgcttg cagagatctc ctgctggatg tggaccggcg agcggcagtc 840
      gacgcggtg cggatccggt acctggacgc ggcgtgcgcg caggacgtgt cctcttcga 900
      caccgacgtg cgcacctcgc acgtcatcta cgccatcaac gcggacgcgc tgggtgggtgc 960
35      aggacgccat cagcgagaag ctgggcaacc tcatccacta catggccacc ttcgtggcg 1020
      gcttcgtcgt gggttcacc gccgcctggc agctggcgct cgtcacgctc gccgtcgtgc 1080
      cgctcatcgc cgtcatcggg gggtcagcg ccgcccgcgt cgccaagctc tcctccagg 1140
      gccaggacgc gctgtcgggc gccagcggca tcgaggagca ggcgtcgcg cagatacggg 1200
      tcgtgcaggc cttcgtcggc gaggagcgcg aaatgcgggc gtactcggcg gcgttggccg 1260
      tcgcgagaa gatcggtac cgcagcggct tcgccaaggg gctcggcctc ggcggcacct 1320
40      acttcaccgt cttctgctg tacggcctc tgccttggtg cggcgagac ctcgtccgcg 1380
      gaaccacacc aacggagggc tcgccatcgc caccatgttc tccgtcatga tcggcgggct 1440
      gtaagatgat cagtttctcc cgggtctctc tgttcttcgc tcatgacaca gcatgtacta 1500
      cgtacgctta ctggtctgtg tctgtgtgtg tgtggatcgc ctgctccag gccctcggg 1560
      cagtcggcgc cagcatggc cgcgttcgcc aaggcgcgcg tggcgccgc caagatcttc 1620
45      cgcatcatcg accacaggcc ggcatctcc tcgcgggacg gcgaggacgg cggcgccgtg 1680
      gagctggagt cggtagcggg gcgggtggag atgagggcg tggacttcgc gtaccctcgc 1740
      cggccggacg tcccatacct gcgcggcttc tcgctcagcg tgcccgcgg caagaccatc 1800
      gcgctggtgg gcagctccgg ctccgggaag agcacgggtg tgcgtcctc cgagagggtc 1860
      tacgaccca gcgcaggtat acatagtag ctaccaatc tagctttagc gcattgatta 1920
50      attagtgttg gagtacatt gcttgccaat tgccattgcc atcacacatc agcagctacc 1980
      atacattgcc aactgccatt gctgctgctc tgcgtgggtg ttagtagggg aagaagcttc 2040
      cactgtagca ggagtacatt gcaaacagga agtgaatttt gcacgtggga aatgaagaag 2100
      tgaatgcttg gagcagagct ggcggcctc atgggctgct tacctactat ctagtcaacc 2160
      aagcatccct gttcttctc tgttatggt caaggcatc acaccagctt agaaacttag 2220
55      aaagaagcta agccctttgt tttaattttc ctaaaaaaa atttttggtt aaaatttttt 2280

```

60

65

ES 2 288 881 T3

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60

```

taaagtttcc cttgtccaca atcccaaatt cttttaagaa gtattaatat atgacgaaaa 2340
ttaaaaccta attgcaccag ttttggtcga aattgaccga aaacgatctt ttgagcctaa 2400
ttagtcccat gattggacaa tatttgtcaa atacaaacga aagtgtctgcc gtatcgattt 2460
tgcaaaaggtt ttcggaacta aacaaagcct ggtgcctgca acgcgagaca aagaaaacta 2520
tttgctgggc aaagtgccac tattgcacat gcatgccact ctttgagcct tgaccgactg 2580
actgactact cagagtagga gtggttcaat tgtattgaca tgtagtagga gtactcgtat 2640
gctatagtag tcctgtagct ttttcaaaca aaaaaaaaag agaaagaaag aaatgaagtc 2700
tgaaatttgt tggttttggc agggcaaatc ttgctggacg ggeatgatct caagtgcgtg 2760
aagctccggt ggtccggca gcagattggt ctggtgagcc aggagccgac gctgttcgcg 2820
acgagcatca aggagaacct gctgctgggg cgggacagtc agagtgcgac gcaggccgag 2880
atggaggagg gccccagggt ggccaacgcg cactccttca tcgtcaagct ccccgacggc 2940
tacgacacgc aggtccgtcc cgtatagcta gctcactagc tgcactgcca cttctctcgc 3000
ttgctcccca ccgttgctgc ctggttctct ccaatccact tgtcgtgtgc tggaccacac 3060
gtgcctgctt gcctagctgc tccacatctg ctttccctgt ccaaccttat gcaactcact 3120
ctaatactat atcaaataca tttctagagt ttaaagctta tcttagaata aatgcatctt 3180
tagtacgag acaacctaac ttcagttggt gttgttgttt tttttacttt ctctcttctc 3240
acaataacta tgattacgtc tttacagcga tcttttttat tccaaaccta aaaatgcatg 3300
cactcactct aaaagcgcaa agggagcatc tttttttccc ccatcatctg cacgcagcct 3360
tttcttttcc tcatgtcacg agggactgaa ggtgtgtatg cagcgtcaag tcatccatcc 3420
gttccacttg gattaggacg ggggggcca tttttaggcc cettgttgcc attcgcattt 3480
tgttgttgtg ggtttgcca caagaaactt gccaggttgc ttttgttatc acgcacagga 3540
caggagaggt ctttttctcg acacaagctc tacagcctct actaaactag cacttgctga 3600
tgagagcaga ggatgaatgg acgatgaaca tctagagtga gagagaaaaa aatgttaata 3660
ataataaaaa gtatagtagc gattaagaat caacctgggg tacgtaggaa gaggtacaat 3720
ccctaggaat ctatagtagt agaagtagtg gaggagttgg gggagtggaa cggaacaaat 3780
tccgagttgg tattttgccg ggaatgtcaa gttgattttt gatcctagtg caagcaagaa 3840
ttatcaatca ctcagactca gcctgtctgt gtctgtccac cccagctctt gctactctac 3900
ttactactgt gctactagtg gtaggtagg tatcttacct aaactgttat tataaactgt 3960
catctgagaa agagagccag tcaaaccctat gctgctgctt attttaatca ctgtcaaatg 4020
gcaggcaggc aggcagctct gttagttaat aacatctggg aagggtttaa tcaaaccaaa 4080
tcaaatacaga cgaaatctag aggccacatg ggatggggcc atatgtactg tactagcata 4140
actagcggct agattttatt agaacacgga ctcaactcc cataactata actgacttga 4200
tcatgattcc ttgccaagca atgctcgcat gcccatgcat gcatcatccc tgggtcaaat 4260
caaacactct ccaccgtcag ggaataagac ttattatttt attaacaatt caatttttat 4320
ttattaatta cgtctggacg aggagtactg gtttatttga tgagagacat ggcagtccaa 4380
gtcaaatctg tttgtctgac ctggcggtg atggccggtg caggttgggg agcgcggcct 4440
gcagctctcc ggtgggcaga agcagcgcat cgccatcgcc cgcgcatgac tcaagaaccc 4500
cgccatctct ctgctggacg aggccaccag cgcgctggac tccgagtctg agaagctcgt 4560
gcaggaggcg ctggaccgct tcatgatcgg gcgcaccacc ctggtgatcg cgcacaggat 4620
gtccaccatc cgcaaggccg acgtggtggc cgtgctgcag ggcggccccc tctccgagat 4680
gggcgcgcac gacgagttga tggccaaggg cgagaacggc acttacgcca agttcatccg 4740
catgcaggag caggcgcacg aggcgcggtt cgtcaacgcc cgccgcagca gcgccaggcc 4800
ctccagcgcc cgcaactccg tcagctcgcc catcatgacg cgcaactcct cctacggccg 4860
ctccccatac tcccgcgcgc tctccgactt ctccacctcc gacttcaccc tctccatcca 4920
cgaccgcgac caccaccacc ggacgatggc cgacaagcag ctgcggttcc gcgcggcgcc 4980
cagctccttc ctccgctcgc ccagatgaa ctcgcccgag tgggcctacg cgctcgtcgg 5040
ctccctgggc tccatggtct gcggtcctt cagcgccatc ttcgccctaca tctcagcgc 5100
cgtgctcagc gtctactacg cgccggaccc tcgctacatg aagcgcgaga tcgccaagta 5160
ctgctactct ctcatcgcca tgtcctccgc ggcgctgctg ttcaacacgg tgagcacgt 5220
gttctgggac acggtcggcg agaacctcac gaagcgtgtg cgcgagaaga tgttcgccc 5280
cgtgctccgc aacgagatcg cctggttcga cgccgacgag aacgccagcg cgcgctcgc 5340
catgcaggcg gcgctcgacg cccagaacct gcgctccgcc atcggggacc gtatctccgt 5400
catcgtccag aactcggcgc tcatgctcgt gcgctgcacc gcgggcttcg tcttccagt 5460
gcgctcgcg ctctgctcct tcgctcgtt cccgctcgtc gtggccgcca ccgtgctgca 5520
gaagatgttc atgaagggt tctcggggga cctggaggcc gcgcacgcca gggccacgca 5580
gatcgcgggc gaggcggtg ccaacctgcg caccgtggcc gcgttcaacg cggagcgcaa 5640
gatcacgggg ctgttcgagg ccaacctgcg cggcccgctc cggcgctgct tctggaagg 5700
gcagatcgcc ggcagcggt acggcggtgc gcagttcctg ctgtacgct cctacgcgt 5760
ggggctgtgg tacgcgcgct ggctggtgaa gcacggcgctg tccgacttct cgcgcacat 5820
ccgcgtgttc atggtgctga tgggtctccg caacggcgcc gccgagacgc tgacgctggc 5880
gccggacttt gtcaaggcg ggcgcgcat gcggtccgtg ttcgagacca tcgaccggaa 5940

```

ES 2 288 881 T3

```

5      aacggaggtg gagcccgacg acgtggacgc ggcgccggtg ccggagcggc ccaagggcga 6000
      ggtggagctg aagcacgtgg acttctcgtg cccgtcgcgg ccggacatcc aggtgttccg 6060
      cgacctgagc ctccggggcg gcgccgggaa gacgctggcg ctggtgggtc cgagcgggtg 6120
      cggaagagc tcggtgctgg cgctggtgca gcggttctac gagccacgt ccgggcgcgt 6180
      gctcctggac ggcaaggacg tgcgcaagta caacctgcgg gcgctgcggc scgtggtggc 6240
      ggtggcgccg caggagccgt tcctgttcgc ggcgagcatc cacgacaaca tcgcgtacgg 6300
      gcgcgagggc gcgacggagg cggaggtggt ggaggcggcg acgcaggcga acgcgcaccg 6360
      gttcatcgcg gcgctgccgg agggctacgg gacgcagggt ggcgagcgcg gggtgcagct 6420
      gtcggggcgg cgagcgcgagc gcatcgcgat cgcgcgcgcg ctggtgaagc aggcggccat 6480
10     cgtgctgctg gacgagggca ccagcgcgct ggacgccgag tcggagcggc gcgtgcagga 6540
      ggcgctggag cgcgcggggt ccgggcgcac caccatcgtg gtggcgacc ggctggccac 6600
      ggtgcgcggc gcgcacacca tcgcggtcat cgacgacggc aaggtggcgg agcaggggtc 6660
      gcactcgcac ctgctcaagc accatcccga cgggtgctac gcgcggatgc tgcagctgca 6720
      gcggctgacg ggcgggtgcc gcgccgggc cgcgcgcgtc gtcgtccaac ggggcggcgg 6780
15     cgtaggatgg atggatggat catggatgag tttgttcc tgaataa 6827

```

```

<210> 8
20    <211> 4213
      <212> ADN
      <213> Sorgo bicolor
      <220>
25    <221> CDS
      <222> (1)..(4206)

      <400> 8
30

```

```

      atg tct acc aac gac ccg gac gag atc agg gcg cgc gtc gtc gtc ctc 48
      Met Ser Thr Asn Asp Pro Asp Glu Ile Arg Ala Arg Val Val Val Leu
      1 5 10 15
35     ggc gcc cct cat gcc gac gac gac gcc ggc gac gag tgg gcc cgc ccc 96
      Gly Ala Pro His Ala Asp Asp Asp Ala Gly Asp Glu Trp Ala Arg Pro
      20 25 30
40     gag ctc gag gcc ttc cac ctc ccc tct ccc gcc cac cag cct cct ggc 144
      Glu Leu Glu Ala Phe His Leu Pro Ser Pro Ala His Gln Pro Pro Gly
      35 40 45
      ttc cac cta gcc gct ggg cac caa ccg gaa gct gca gca gag caa ccc 192
      Phe His Leu Ala Ala Gly His Gln Pro Glu Ala Ala Ala Glu Gln Pro
      50 55 60
45     acc acg ctc cct gct gcc cgc cgc acc agc gac aca tcc act gct gct 240
      Thr Thr Leu Pro Ala Ala Arg Arg Thr Ser Asp Thr Ser Thr Ala Ala
      65 70 75 80
50     ggt gct gct cct cct tct cct tcg ccg cct ccg ccg ccg gct cct ttg 288
      Gly Ala Ala Pro Pro Ser Pro Ser Pro Pro Pro Pro Pro Ala Pro Leu
      85 90 95
      gag atg gac cag ccg ccc aat gcc aag ccg gcc tcc tcc tcc gcc gcc 336
      Glu Met Asp Gln Pro Pro Asn Ala Lys Pro Ala Ser Ser Ser Ala Ala
      100 105 110
55     gcc gcc ggc gcc aat gac aac aag aag ccc acc ccg ccc gcc gcg ctg 384
      Ala Ala Gly Ala Asn Asp Asn Lys Lys Pro Thr Pro Pro Ala Ala Leu
      115 120 125
60     cgc gac ctc ttc cgc ttc gcc gac ggc ctc gac tgc gcg ctc atg ctc 432
      Arg Asp Leu Phe Arg Phe Ala Asp Gly Leu Asp Cys Ala Leu Met Leu

```

ES 2 288 881 T3

	130	135	140	
5	gtc ggc acg ctc ggc ggc ctc gtc cac ggc tgc tgc ctc ccc gtc ttc 480 Val Gly Thr Leu Gly Ala Leu Val His Gly Cys Ser Leu Pro Val Phe 145 150 155 160			
	ctc cgc ttc ttc gcc gac ctc gtc gac tcc ttc ggc tcc cac gcc aac 528 Leu Arg Phe Phe Ala Asp Leu Val Asp Ser Phe Gly Ser His Ala Asn 165 170 175			
10	gac ccg gac acc atg gtc cgc ctc gtc gtc aag tac gcc ttc tac ttc 576 Asp Pro Asp Thr Met Val Arg Leu Val Val Lys Tyr Ala Phe Tyr Phe 180 185 190			
15	ctc gtc gtc gga gcc gca atc tgg ggc tcc tca tgg gca gag atc ttc 624 Leu Val Val Gly Ala Ala Ile Trp Ala Ser Ser Trp Ala Glu Ile Ser 195 200 205			
	tgc tgg atg tgg acc ggc gag cgg cag tgc acg cgg atg cgg atc cgg 672 Cys Trp Met Trp Thr Gly Glu Arg Gln Ser Thr Arg Met Arg Ile Arg 210 215 220			
20	tac ctg gac ggc ggc ctg cgg cag gac gtg tcc ttc ttc gac acc gac 720 Tyr Leu Asp Ala Ala Leu Arg Gln Asp Val Ser Phe Phe Asp Thr Asp 225 230 235 240			
25	gtg cgc acc tgc gac gtc atc tac gcc atc aac gcg gac gcc gtg gtg 768 Val Arg Thr Ser Asp Val Ile Tyr Ala Ile Asn Ala Asp Ala Val Val 245 250 255			
	ggt gca gga cgc cat cag cga gaa gct ggg caa cct cat cca cta cat 816 Gly Ala Gly Arg His Gln Arg Glu Ala Gly Gln Pro His Pro Leu His 260 265 270			
30	ggc cac ctt cgt ggc ggc ctt cgt cgt ggg ctt cac cgc cgc ctg gca 864 Gly His Leu Arg Gly Gly Leu Arg Arg Gly Leu His Arg Arg Leu Ala 275 280 285			
35	gct ggc gct cgt cac gct cgc cgt cgt gcc gct cat cgc cgt cat cgg 912 Ala Gly Ala Arg His Ala Arg Arg Arg Ala Ala His Arg Arg His Arg 290 295 300			
	ggg gct cag cgc cgc cgc gct cgc caa gct ctc ctc cag gag cca gga 960 Gly Ala Gln Arg Arg Arg Ala Arg Gln Ala Leu Leu Gln Glu Pro Gly 305 310 315 320			
40	cgc gct gtc ggg cgc cag cgg cat cgc gga gca ggc gct cgc gca gat 1008 Arg Ala Val Gly Arg Gln Arg His Arg Gly Ala Gly Ala Arg Ala Asp 325 330 335			
45	acg gat cgt gca ggc ctt cgt cgg cga gga gcg cga aat gcg ggc gta 1056 Thr Asp Arg Ala Gly Leu Arg Arg Arg Gly Ala Arg Asn Ala Gly Val 340 345 350			
	ctc ggc ggc gtt ggc cgt cgc gca gaa gat cgg cta ccg cag cgg ctt 1104 Leu Gly Gly Val Gly Arg Arg Ala Glu Asp Arg Leu Pro Gln Arg Leu 355 360 365			
50	cgc caa ggg gct cgg cct cgg cgg cac cta ctt cac cgt ctt ctg ctg 1152 Arg Gln Gly Ala Arg Pro Arg Arg His Leu Leu His Arg Leu Leu Leu 370 375 380			

ES 2 288 881 T3

	cta	cgg	cct	cct	gct	ctg	gta	cgg	cgg	aca	cct	cgt	ccg	cgg	aac	cac	1200
	Leu	Arg	Pro	Pro	Ala	Leu	Val	Arg	Arg	Thr	Pro	Arg	Pro	Arg	Asn	His	
	385					390					395				400		
5	acc	aac	gga	ggg	ctc	gcc	atc	gcc	acc	atg	ttc	tcc	gtc	atg	atc	ggc	1248
	Thr	Asn	Gly	Gly	Leu	Ala	Ile	Ala	Thr	Met	Phe	Ser	Val	Met	Ile	Gly	
					405					410					415		
10	ggg	ctg	gcc	ctc	ggg	cag	tcg	gcg	ccg	agc	atg	gcc	gcg	ttc	gcc	aag	1296
	Gly	Leu	Ala	Leu	Gly	Gln	Ser	Ala	Pro	Ser	Met	Ala	Ala	Phe	Ala	Lys	
					420				425						430		
15	gcg	cgc	gtg	gcg	gcc	gcc	aag	atc	ttc	cgc	atc	atc	gac	cac	agg	ccg	1344
	Ala	Arg	Val	Ala	Ala	Ala	Lys	Ile	Phe	Arg	Ile	Ile	Asp	His	Arg	Pro	
					435			440						445			
20	ggc	atc	tcc	tcg	cgg	gac	ggc	gag	gac	ggc	ggc	ggc	gtg	gag	ctg	gag	1392
	Gly	Ile	Ser	Ser	Arg	Asp	Gly	Glu	Asp	Gly	Gly	Gly	Val	Glu	Leu	Glu	
		450					455					460					
25	tcg	gtg	acg	ggg	cgg	gtg	gag	atg	agg	ggc	gtg	gac	ttc	gcg	tac	ccg	1440
	Ser	Val	Thr	Gly	Arg	Val	Glu	Met	Arg	Gly	Val	Asp	Phe	Ala	Tyr	Pro	
		465				470					475					480	
30	tcg	cgg	ccg	gac	gtc	ccc	atc	ctg	cgc	ggc	ttc	tcg	ctc	agc	gtg	ccc	1488
	Ser	Arg	Pro	Asp	Val	Pro	Ile	Leu	Arg	Gly	Phe	Ser	Leu	Ser	Val	Pro	
					485					490					495		
35	gcc	ggc	aag	acc	atc	gcg	ctg	gtg	ggc	agc	tcc	ggc	tcc	ggg	aag	agc	1536
	Ala	Gly	Lys	Thr	Ile	Ala	Leu	Val	Gly	Ser	Ser	Gly	Ser	Gly	Lys	Ser	
					500				505					510			
40	acg	gtg	gtg	tcg	ctc	ctc	gag	agg	ttc	tac	gac	ccc	agc	gca	ggg	caa	1584
	Thr	Val	Val	Ser	Leu	Leu	Glu	Arg	Phe	Tyr	Asp	Pro	Ser	Ala	Gly	Gln	
			515				520						525				
45	atc	ttg	ctg	gac	ggg	cat	gat	ctc	aag	tcg	ctg	aag	ctc	cgg	tgg	ctc	1632
	Ile	Leu	Leu	Asp	Gly	His	Asp	Leu	Lys	Ser	Leu	Lys	Leu	Arg	Trp	Leu	
		530					535					540					
50	cgg	cag	cag	att	ggt	ctg	gtg	agc	cag	gag	ccg	acg	ctg	ttc	gcg	acg	1680
	Arg	Gln	Gln	Ile	Gly	Leu	Val	Ser	Gln	Glu	Pro	Thr	Leu	Phe	Ala	Thr	
		545				550					555				560		
55	agc	atc	aag	gag	aac	ctg	ctg	ctg	ggg	cgg	gac	agt	cag	agt	gcg	acg	1728
	Ser	Ile	Lys	Glu	Asn	Leu	Leu	Leu	Gly	Arg	Asp	Ser	Gln	Ser	Ala	Thr	
					565					570					575		
60	cag	gcc	gag	atg	gag	gag	gcc	gcc	agg	gtg	gcc	aac	gcg	cac	tcc	ttc	1776
	Gln	Ala	Glu	Met	Glu	Glu	Ala	Ala	Arg	Val	Ala	Asn	Ala	His	Ser	Phe	
					580				585					590			
65	atc	gtc	aag	ctc	ccc	gac	ggc	tac	gac	acg	cag	gtt	ggg	gag	cgc	ggc	1824
	Ile	Val	Lys	Leu	Pro	Asp	Gly	Tyr	Asp	Thr	Gln	Val	Gly	Glu	Arg	Gly	
			595				600						605				
70	ctg	cag	ctc	tcc	ggt	ggg	cag	aag	cag	cgc	atc	gcc	atc	gcc	cgc	gcc	1872
	Leu	Gln	Leu	Ser	Gly	Gly	Gln	Lys	Gln	Arg	Ile	Ala	Ile	Ala	Arg	Ala	
			610				615					620					

ES 2 288 881 T3

	atg ctc aag aac ccc gcc atc ctg ctg ctg gac gag gcc acc agc gcg	1920
	Met Leu Lys Asn Pro Ala Ile Leu Leu Leu Asp Glu Ala Thr Ser Ala	
	625 630 635 640	
5	ctg gac tcc gag tct gag aag ctc gtg cag gag gcg ctg gac cgc ttc	1968
	Leu Asp Ser Glu Ser Glu Lys Leu Val Gln Glu Ala Leu Asp Arg Phe	
	645 650 655	
10	atg atc ggg cgc acc acc ctg gtg atc gcg cac agg atg tcc acc atc	2016
	Met Ile Gly Arg Thr Thr Leu Val Ile Ala His Arg Met Ser Thr Ile	
	660 665 670	
	cgc aag gcc gac gtg gtg gcc gtg ctg cag gcc gcc ccc gtc tcc gag	2064
	Arg Lys Ala Asp Val Val Ala Val Leu Gln Gly Gly Pro Val Ser Glu	
	675 680 685	
15	atg ggc gcg cac gac gag ttg atg gcc aag gcc gag aac gcc act tac	2112
	Met Gly Ala His Asp Glu Leu Met Ala Lys Gly Glu Asn Gly Thr Tyr	
	690 695 700	
20	gcc aag ttc atc cgc atg cag gag cag gcg cac gag gcg gcg ttc gtc	2160
	Ala Lys Phe Ile Arg Met Gln Glu Gln Ala His Glu Ala Ala Phe Val	
	705 710 715 720	
	aac gcc cgc cgc agc agc gcc agg ccc tcc agc gcc cgc aac tcc gtc	2208
	Asn Ala Arg Arg Ser Ser Ala Arg Pro Ser Ser Ala Arg Asn Ser Val	
	725 730 735	
25	agc tcg ccc atc atg acg cgc aac tcc tcc tac gcc cgc tcc cca tac	2256
	Ser Ser Pro Ile Met Thr Arg Asn Ser Ser Tyr Gly Arg Ser Pro Tyr	
	740 745 750	
30	tcc cgc cgc ctc tcc gac ttc tcc acc tcc gac ttc acc ctc tcc atc	2304
	Ser Arg Arg Leu Ser Asp Phe Ser Thr Ser Asp Phe Thr Leu Ser Ile	
	755 760 765	
	cac gac ccg cac cac cac cac cgg acg atg gcc gac aag cag ctc gcg	2352
	His Asp Pro His His His His Arg Thr Met Ala Asp Lys Gln Leu Ala	
	770 775 780	
35	ttc cgc gcc gcc gcc agc tcc ttc ctc cgc ctc gcc agg atg aac tcg	2400
	Phe Arg Ala Gly Ala Ser Ser Phe Leu Arg Leu Ala Arg Met Asn Ser	
	785 790 795 800	
40	ccc gag tgg gcc tac gcg ctc gtc gcc tcc ctg gcc tcc atg gtc tgc	2448
	Pro Glu Trp Ala Tyr Ala Leu Val Gly Ser Leu Gly Ser Met Val Cys	
	805 810 815	
	ggc tcc ttc agc gcc atc ttc gcc tac atc ctc agc gcc gtg ctc agc	2496
	Gly Ser Phe Ser Ala Ile Phe Ala Tyr Ile Leu Ser Ala Val Leu Ser	
	820 825 830	
45	gtc tac tac gcg ccg gac cct cgc tac atg aag cgc gag atc gcc aag	2544
	Val Tyr Tyr Ala Pro Asp Pro Arg Tyr Met Lys Arg Glu Ile Ala Lys	
	835 840 845	
50	tac tgc tac ctg ctc atc gcc atg tcc tcc gcg gcg ctg ctg ttc aac	2592
	Tyr Cys Tyr Leu Leu Ile Gly Met Ser Ser Ala Ala Leu Leu Phe Asn	
	850 855 860	
55	acg gtg cag cac gtg ttc tgg gac acg gtc gcc gag aac ctc acg aag	2640

ES 2 288 881 T3

	Thr	Val	Gln	His	Val	Phe	Trp	Asp	Thr	Val	Gly	Glu	Asn	Leu	Thr	Lys	
	865					870					875					880	
5	cgt	gtg	cgc	gag	aag	atg	ttc	gcc	gcc	gtg	ctc	cgc	aac	gag	atc	gcc	2688
	Arg	Val	Arg	Glu	Lys	Met	Phe	Ala	Ala	Val	Leu	Arg	Asn	Glu	Ile	Ala	
				885						890					895		
	tgg	ttc	gac	gcc	gac	gag	aac	gcc	agc	gcg	cgc	gtc	gcc	gcc	agg	ctc	2736
	Trp	Phe	Asp	Ala	Asp	Glu	Asn	Ala	Ser	Ala	Arg	Val	Ala	Ala	Arg	Leu	
				900					905						910		
10	gcg	ctc	gac	gcc	cag	aac	gtg	cgc	tcc	gcc	atc	ggg	gac	cgt	atc	tcc	2784
	Ala	Leu		Ala	Gln	Asn	Val	Arg	Ser	Ala	Ile	Gly	Asp	Arg	Ile	Ser	
				915				920					925				
15	gtc	atc	gtc	cag	aac	tcg	gcg	ctc	atg	ctc	gtc	gcc	tgc	acc	gcg	ggc	2832
	Val	Ile	Val	Gln	Asn	Ser	Ala	Leu	Met	Leu	Val	Ala	Cys	Thr	Ala	Gly	
				930				935					940				
	ttc	gtc	ctc	cag	tgg	cgc	ctc	gcg	ctc	gtg	ctc	ctc	gcc	gtc	ttc	ccg	2880
	Phe	Val	Leu	Gln	Trp	Arg	Leu	Ala	Leu	Val	Leu	Leu	Ala	Val	Phe	Pro	
						950					955					960	
20	ctc	gtc	gtg	gcc	gcc	acc	gtg	ctg	cag	aag	atg	ttc	atg	aag	ggc	ttc	2928
	Leu	Val	Val	Ala	Ala	Thr	Val	Leu	Gln	Lys	Met	Phe	Met	Lys	Gly	Phe	
						965				970					975		
25	tcg	ggg	gac	ctg	gag	gcc	gcg	cac	gcc	agg	gcc	acg	cag	atc	gcg	ggc	2976
	Ser	Gly	Asp	Leu	Glu	Ala	Ala	His	Ala	Arg	Ala	Thr	Gln	Ile	Ala	Gly	
				980					985					990			
	gag	gcc	gtg	gcc	aac	ctg	cgc	acc	gtg	gcc	gcg	ttc	aac	gcg	gag	cgc	3024
	Glu	Ala	Val	Ala	Asn	Leu	Arg	Thr	Val	Ala	Ala	Phe	Asn	Ala	Glu	Arg	
				995				1000					1005				
30	aag	atc	acg	ggg	ctg	ttc	gag	gcc	aac	ctg	cgc	ggc	ccg	ctc	cgg	cgc	3072
	Lys	Ile	Thr	Gly	Leu	Phe	Glu	Ala	Asn	Leu	Arg	Gly	Pro	Leu	Arg	Arg	
		1010					1015					1020					
35	tgc	ttc	tgg	aag	ggg	cag	atc	gcc	ggc	agc	ggc	tac	ggc	gtg	gcg	cag	3120
	Cys	Phe	Trp	Lys	Gly	Gln	Ile	Ala	Gly	Ser	Gly	Tyr	Gly	Val	Ala	Gln	
						1030					1035					1040	
	ttc	ctg	ctg	tac	gcg	tcc	tac	gcg	ctg	ggg	ctg	tgg	tac	gcg	gcg	tgg	3168
	Phe	Leu	Leu	Tyr	Ala	Ser	Tyr	Ala	Leu	Gly	Leu	Trp	Tyr	Ala	Ala	Trp	
					1045					1050					1055		
40	ctg	gtg	aag	cac	ggc	gtg	tcc	gac	ttc	tcg	cgc	acc	atc	cgc	gtg	ttc	3216
	Leu	Val	Lys	His	Gly	Val	Ser	Asp	Phe	Ser	Arg	Thr	Ile	Arg	Val	Phe	
				1060				1065						1070			
45	atg	gtg	ctg	atg	gtg	tcc	gcc	aac	ggc	gcc	gcc	gag	acg	ctg	acg	ctg	3264
	Met	Val	Leu	Met	Val	Ser	Ala	Asn	Gly	Ala	Ala	Glu	Thr	Leu	Thr	Leu	
				1075				1080					1085				
	gcg	ccg	gac	ttt	gtc	aag	ggc	ggg	cgc	gcg	atg	cgg	tcc	gtg	ttc	gag	3312
	Ala	Pro	Asp	Phe	Val	Lys	Gly	Gly	Arg	Ala	Met	Arg	Ser	Val	Phe	Glu	
				1090				1095				1100					
50	acc	atc	gac	cgg	aaa	acg	gag	gtg	gag	ccc	gac	gac	gtg	gac	gcg	gcg	3360
	Thr	Ile	Asp	Arg	Lys	Thr	Glu	Val	Glu	Pro	Asp	Asp	Val	Asp	Ala	Ala	

ES 2 288 881 T3

	1105	1110	1115	1120	
5	ccg gtg ccg gag ccg ccc aag ggc gag gtg gag ctg aag cac gtg gac Pro Val Pro Glu Arg Pro Lys Gly Glu Val Glu Leu Lys His Val Asp	3408			
	1125	1130	1135		
10	ttc tcg tac ccg tcg ccg ccg gac atc cag gtg ttc cgc gac ctg agc Phe Ser Tyr Pro Ser Arg Pro Asp Ile Gln Val Phe Arg Asp Leu Ser	3456			
	1140	1145	1150		
15	ctc ccg gcg cgc gcc ggg aag acg ctg gcg ctg gtg ggt ccg agc ggg Leu Arg Ala Arg Ala Gly Lys Thr Leu Ala Leu Val Gly Pro Ser Gly	3504			
	1155	1160	1165		
20	tgc ggc aag agc tcg gtg ctg gcg ctg gtg cag ccg ttc tac gag ccc Cys Gly Lys Ser Ser Val Leu Ala Leu Val Gln Arg Phe Tyr Glu Pro	3552			
	1170	1175	1180		
25	acg tcc ggg cgc gtg ctc ctg gac ggc aag gac gtg cgc aag tac aac Thr Ser Gly Arg Val Leu Leu Asp Gly Lys Asp Val Arg Lys Tyr Asn	3600			
	1185	1190	1195	1200	
30	ctg ccg gcg ctg ccg cgc gtg gtg gcg gtg gcg ccg cag gag ccg ttc Leu Arg Ala Leu Arg Arg Val Val Ala Val Ala Pro Gln Glu Pro Phe	3648			
	1205	1210	1215		
35	ctg ttc gcg gcg agc atc cac gac aac atc gcg tac ggg cgc gag ggc Leu Phe Ala Ala Ser Ile His Asp Asn Ile Ala Tyr Gly Arg Glu Gly	3696			
	1220	1225	1230		
40	gcg acg gag gcg gag gtg gtg gag gcg gcg acg cag gcg aac gcg cac Ala Thr Glu Ala Glu Val Val Glu Ala Ala Thr Gln Ala Asn Ala His	3744			
	1235	1240	1245		
45	ccg ttc atc gcg gcg ctg ccg gag ggc tac ggg acg cag gtg ggc gag Arg Phe Ile Ala Ala Leu Pro Glu Gly Tyr Gly Thr Gln Val Gly Glu	3792			
	1250	1255	1260		
50	cgc ggg gtg cag ctg tcg ggc ggg cag ccg cag ccg atc gcg atc gcg Arg Gly Val Gln Leu Ser Gly Gly Gln Arg Gln Arg Ile Ala Ile Ala	3840			
	1265	1270	1275	1280	
55	cgc gcg ctg gtg aag cag gcg gcc atc gtg ctg ctg gac gag gcg acc Arg Ala Leu Val Lys Gln Ala Ala Ile Val Leu Leu Asp Glu Ala Thr	3888			
	1285	1290	1295		
60	agc gcg ctg gac gcc gag tcg gag ccg tgc gtg cag gag gcg ctg gag Ser Ala Leu Asp Ala Glu Ser Arg Cys Val Gln Glu Ala Leu Glu	3936			
	1300	1305	1310		
65	cgc gcg ggg tcc ggg cgc acc acc atc gtg gtg gcg cac ccg ctg gcc Arg Ala Gly Ser Gly Arg Thr Thr Ile Val Val Ala His Arg Leu Ala	3984			
	1315	1320	1325		
70	acg gtg cgc ggc gcg cac acc atc gcg gtc atc gac gac ggc aag gtg Thr Val Arg Gly Ala His Thr Ile Ala Val Ile Asp Asp Gly Lys Val	4032			
	1330	1335	1340		
75	gcg gag cag ggg tcg cac tcg cac ctg ctc aag cac cat ccc gac ggg Ala Glu Gln Gly Ser His Ser His Leu Leu Lys His His Pro Asp Gly	4080			
	1345	1350	1355	1360	
80	tgc tac gcg ccg atg ctg cag ctg cag ccg ctg acg ggc ggg tgc cgc Cys Tyr Ala Arg Met Leu Gln Leu Gln Arg Leu Thr Gly Gly Cys Arg	4128			
	1365	1370	1375		
85	gcc ccg gcc gcc gcc gtc gtc gtc caa ccg ggc cgc cgc gta gga tgg Ala Arg Ala Ala Val Val Val Gln Arg Gly Arg Arg Val Gly Trp	4176			
	1380	1385	1390		
90	atg gat gga tca tgg atg agt ttg gtt cct tgataaa Met Asp Gly Ser Trp Met Ser Leu Val Pro	4213			
	1395	1400			

ES 2 288 881 T3

<210> 9

<211> 1402

<212> PRT

5 <213> Sorgo bicolor

<400> 9

10	Met	Ser	Thr	Asn	Asp	Pro	Asp	Glu	Ile	Arg	Ala	Arg	Val	Val	Val	Leu
	1				5					10					15	
	Gly	Ala	Pro	His	Ala	Asp	Asp	Asp	Ala	Gly	Asp	Glu	Trp	Ala	Arg	Pro
				20					25					30		
15	Glu	Leu	Glu	Ala	Phe	His	Leu	Pro	Ser	Pro	Ala	His	Gln	Pro	Pro	Gly
			35					40					45			
	Phe	His	Leu	Ala	Ala	Gly	His	Gln	Pro	Glu	Ala	Ala	Ala	Glu	Gln	Pro
		50					55						60			
20	Thr	Thr	Leu	Pro	Ala	Ala	Arg	Arg	Thr	Ser	Asp	Thr	Ser	Thr	Ala	Ala
	65					70					75					80
	Gly	Ala	Ala	Pro	Pro	Ser	Pro	Ser	Pro	Pro	Pro	Pro	Pro	Ala	Pro	Leu
25					85					90					95	
	Glu	Met	Asp	Gln	Pro	Pro	Asn	Ala	Lys	Pro	Ala	Ser	Ser	Ser	Ala	Ala
				100					105					110		
	Ala	Ala	Gly	Ala	Asn	Asp	Asn	Lys	Lys	Pro	Thr	Pro	Pro	Ala	Ala	Leu
30			115					120					125			
	Arg	Asp	Leu	Phe	Arg	Phe	Ala	Asp	Gly	Leu	Asp	Cys	Ala	Leu	Met	Leu
		130					135					140				
35	Val	Gly	Thr	Leu	Gly	Ala	Leu	Val	His	Gly	Cys	Ser	Leu	Pro	Val	Phe
	145					150					155					160
	Leu	Arg	Phe	Phe	Ala	Asp	Leu	Val	Asp	Ser	Phe	Gly	Ser	His	Ala	Asn
				165					170						175	
40	Asp	Pro	Asp	Thr	Met	Val	Arg	Leu	Val	Val	Lys	Tyr	Ala	Phe	Tyr	Phe
				180					185						190	
	Leu	Val	Val	Gly	Ala	Ala	Ile	Trp	Ala	Ser	Ser	Trp	Ala	Glu	Ile	Ser
			195					200					205			
45	Cys	Trp	Met	Trp	Thr	Gly	Glu	Arg	Gln	Ser	Thr	Arg	Met	Arg	Ile	Arg
		210					215					220				

50

55

60

65

ES 2 288 881 T3

	Tyr	Leu	Asp	Ala	Ala	Leu	Arg	Gln	Asp	Val	Ser	Phe	Phe	Asp	Thr	Asp	225	230	235	240
5	Val	Arg	Thr	Ser	Asp	Val	Ile	Tyr	Ala	Ile	Asn	Ala	Asp	Ala	Val	Val	245	250	255	
	Gly	Ala	Gly	Arg	His	Gln	Arg	Glu	Ala	Gly	Gln	Pro	His	Pro	Leu	His	260	265	270	
10	Gly	His	Leu	Arg	Gly	Gly	Leu	Arg	Arg	Gly	Leu	His	Arg	Arg	Leu	Ala	275	280	285	
15	Ala	Gly	Ala	Arg	His	Ala	Arg	Arg	Arg	Ala	Ala	His	Arg	Arg	His	Arg	290	295	300	
	Gly	Ala	Gln	Arg	Arg	Arg	Ala	Arg	Gln	Ala	Leu	Leu	Gln	Glu	Pro	Gly	305	310	315	320
20	Arg	Ala	Val	Gly	Arg	Gln	Arg	His	Arg	Gly	Ala	Gly	Ala	Arg	Ala	Asp	325	330	335	
	Thr	Asp	Arg	Ala	Gly	Leu	Arg	Arg	Arg	Gly	Ala	Arg	Asn	Ala	Gly	Val	340	345	350	
25	Leu	Gly	Gly	Val	Gly	Arg	Arg	Ala	Glu	Asp	Arg	Leu	Pro	Gln	Arg	Leu	355	360	365	
30	Arg	Gln	Gly	Ala	Arg	Pro	Arg	Arg	His	Leu	Leu	His	Arg	Leu	Leu	Leu	370	375	380	
	Leu	Arg	Pro	Pro	Ala	Leu	Val	Arg	Arg	Thr	Pro	Arg	Pro	Arg	Asn	His	385	390	395	400
35	Thr	Asn	Gly	Gly	Leu	Ala	Ile	Ala	Thr	Met	Phe	Ser	Val	Met	Ile	Gly	405	410	415	
	Gly	Leu	Ala	Leu	Gly	Gln	Ser	Ala	Pro	Ser	Met	Ala	Ala	Phe	Ala	Lys	420	425	430	
40	Ala	Arg	Val	Ala	Ala	Ala	Lys	Ile	Phe	Arg	Ile	Ile	Asp	His	Arg	Pro	435	440	445	
45	Gly	Ile	Ser	Ser	Arg	Asp	Gly	Glu	Asp	Gly	Gly	Gly	Val	Glu	Leu	Glu	450	455	460	
	Ser	Val	Thr	Gly	Arg	Val	Glu	Met	Arg	Gly	Val	Asp	Phe	Ala	Tyr	Pro	465	470	475	480
50	Ser	Arg	Pro	Asp	Val	Pro	Ile	Leu	Arg	Gly	Phe	Ser	Leu	Ser	Val	Pro	485	490	495	
	Ala	Gly	Lys	Thr	Ile	Ala	Leu	Val	Gly	Ser	Ser	Gly	Ser	Gly	Lys	Ser	500	505	510	
55	Thr	Val	Val	Ser	Leu	Leu	Glu	Arg	Phe	Tyr	Asp	Pro	Ser	Ala	Gly	Gln	515	520	525	
60	Ile	Leu	Leu	Asp	Gly	His	Asp	Leu	Lys	Ser	Leu	Lys	Leu	Arg	Trp	Leu	530	535	540	

ES 2 288 881 T3

	Arg	Gln	Gln	Ile	Gly	Leu	Val	Ser	Gln	Glu	Pro	Thr	Leu	Phe	Ala	Thr	545	550	555	560
5	Ser	Ile	Lys	Glu	Asn	Leu	Leu	Leu	Gly	Arg	Asp	Ser	Gln	Ser	Ala	Thr	565	570	575	
	Gln	Ala	Glu	Met	Glu	Glu	Ala	Ala	Arg	Val	Ala	Asn	Ala	His	Ser	Phe	580	585	590	
10	Ile	Val	Lys	Leu	Pro	Asp	Gly	Tyr	Asp	Thr	Gln	Val	Gly	Glu	Arg	Gly	595	600	605	
	Leu	Gln	Leu	Ser	Gly	Gly	Gln	Lys	Gln	Arg	Ile	Ala	Ile	Ala	Arg	Ala	610	615	620	
15	Met	Leu	Lys	Asn	Pro	Ala	Ile	Leu	Leu	Leu	Asp	Glu	Ala	Thr	Ser	Ala	625	630	635	640
	Leu	Asp	Ser	Glu	Ser	Glu	Lys	Leu	Val	Gln	Glu	Ala	Leu	Asp	Arg	Phe	645	650	655	
20	Met	Ile	Gly	Arg	Thr	Thr	Leu	Val	Ile	Ala	His	Arg	Met	Ser	Thr	Ile	660	665	670	
	Arg	Lys	Ala	Asp	Val	Val	Ala	Val	Leu	Gln	Gly	Gly	Pro	Val	Ser	Glu	675	680	685	
25	Met	Gly	Ala	His	Asp	Glu	Leu	Met	Ala	Lys	Gly	Glu	Asn	Gly	Thr	Tyr	690	695	700	
	Ala	Lys	Phe	Ile	Arg	Met	Gln	Glu	Gln	Ala	His	Glu	Ala	Ala	Phe	Val	705	710	715	720
30	Asn	Ala	Arg	Arg	Ser	Ser	Ala	Arg	Pro	Ser	Ser	Ala	Arg	Asn	Ser	Val	725	730	735	
	Ser	Ser	Pro	Ile	Met	Thr	Arg	Asn	Ser	Ser	Tyr	Gly	Arg	Ser	Pro	Tyr	740	745	750	
35	Ser	Arg	Arg	Leu	Ser	Asp	Phe	Ser	Thr	Ser	Asp	Phe	Thr	Leu	Ser	Ile	755	760	765	
	His	Asp	Pro	His	His	His	His	Arg	Thr	Met	Ala	Asp	Lys	Gln	Leu	Ala	770	775	780	
40	Phe	Arg	Ala	Gly	Ala	Ser	Ser	Phe	Leu	Arg	Leu	Ala	Arg	Met	Asn	Ser	785	790	795	800
	Pro	Glu	Trp	Ala	Tyr	Ala	Leu	Val	Gly	Ser	Leu	Gly	Ser	Met	Val	Cys	805	810	815	
45	Gly	Ser	Phe	Ser	Ala	Ile	Phe	Ala	Tyr	Ile	Leu	Ser	Ala	Val	Leu	Ser	820	825	830	
	Val	Tyr	Tyr	Ala	Pro	Asp	Pro	Arg	Tyr	Met	Lys	Arg	Glu	Ile	Ala	Lys	835	840	845	
50	Tyr	Cys	Tyr	Leu	Leu	Ile	Gly	Met	Ser	Ser	Ala	Ala	Leu	Leu	Phe	Asn	850	855	860	
55	Thr	Val	Gln	His	Val	Phe	Trp	Asp	Thr	Val	Gly	Glu	Asn	Leu	Thr	Lys				

ES 2 288 881 T3

	865		870		875		880
5	Arg Val Arg Glu Lys Met Phe Ala Ala Val Leu Arg Asn Glu Ile Ala	885		890		895	
	Trp Phe Asp Ala Asp Glu Asn Ala Ser Ala Arg Val Ala Ala Arg Leu	900		905		910	
10	Ala Leu Asp Ala Gln Asn Val Arg Ser Ala Ile Gly Asp Arg Ile Ser	915		920		925	
	Val Ile Val Gln Asn Ser Ala Leu Met Leu Val Ala Cys Thr Ala Gly	930		935		940	
15	Phe Val Leu Gln Trp Arg Leu Ala Leu Val Leu Leu Ala Val Phe Pro	945		950		955	960
	Leu Val Val Ala Ala Thr Val Leu Gln Lys Met Phe Met Lys Gly Phe	965		970		975	
20	Ser Gly Asp Leu Glu Ala Ala His Ala Arg Ala Thr Gln Ile Ala Gly	980		985		990	
	Glu Ala Val Ala Asn Leu Arg Thr Val Ala Ala Phe Asn Ala Glu Arg	995		1000		1005	
25	Lys Ile Thr Gly Leu Phe Glu Ala Asn Leu Arg Gly Pro Leu Arg Arg	1010		1015		1020	
30	Cys Phe Trp Lys Gly Gln Ile Ala Gly Ser Gly Tyr Gly Val Ala Gln	1025		1030		1035	1040
	Phe Leu Leu Tyr Ala Ser Tyr Ala Leu Gly Leu Trp Tyr Ala Ala Trp	1045		1050		1055	
35	Leu Val Lys His Gly Val Ser Asp Phe Ser Arg Thr Ile Arg Val Phe	1060		1065		1070	
	Met Val Leu Met Val Ser Ala Asn Gly Ala Ala Glu Thr Leu Thr Leu	1075		1080		1085	
40	Ala Pro Asp Phe Val Lys Gly Gly Arg Ala Met Arg Ser Val Phe Glu	1090		1095		1100	
45	Thr Ile Asp Arg Lys Thr Glu Val Glu Pro Asp Asp Val Asp Ala Ala	1105		1110		1115	1120
	Pro Val Pro Glu Arg Pro Lys Gly Glu Val Glu Leu Lys His Val Asp	1125		1130		1135	
50	Phe Ser Tyr Pro Ser Arg Pro Asp Ile Gln Val Phe Arg Asp Leu Ser	1140		1145		1150	
	Leu Arg Ala Arg Ala Gly Lys Thr Leu Ala Leu Val Gly Pro Ser Gly	1155		1160		1165	
55	Cys Gly Lys Ser Ser Val Leu Ala Leu Val Gln Arg Phe Tyr Glu Pro	1170		1175		1180	
60	Thr Ser Gly Arg Val Leu Leu Asp Gly Lys Asp Val Arg Lys Tyr Asn	1185		1190		1195	1200

ES 2 288 881 T3

	Leu Arg Ala Leu Arg Arg Val Val Ala Val Ala Pro Gln Glu Pro Phe
	1205 1210 1215
5	Leu Phe Ala Ala Ser Ile His Asp Asn Ile Ala Tyr Gly Arg Glu Gly
	1220 1225 1230
	Ala Thr Glu Ala Glu Val Val Glu Ala Ala Thr Gln Ala Asn Ala His
	1235 1240 1245
10	Arg Phe Ile Ala Ala Leu Pro Glu Gly Tyr Gly Thr Gln Val Gly Glu
	1250 1255 1260
	Arg Gly Val Gln Leu Ser Gly Gly Gln Arg Gln Arg Ile Ala Ile Ala
	1265 1270 1275 1280
15	Arg Ala Leu Val Lys Gln Ala Ala Ile Val Leu Leu Asp Glu Ala Thr
	1285 1290 1295
	Ser Ala Leu Asp Ala Glu Ser Glu Arg Cys Val Gln Glu Ala Leu Glu
	1300 1305 1310
20	Arg Ala Gly Ser Gly Arg Thr Thr Ile Val Val Ala His Arg Leu Ala
	1315 1320 1325
	Thr Val Arg Gly Ala His Thr Ile Ala Val Ile Asp Asp Gly Lys Val
	1330 1335 1340
25	Ala Glu Gln Gly Ser His Ser His Leu Leu Lys His His Pro Asp Gly
	1345 1350 1355 1360
	Cys Tyr Ala Arg Met Leu Gln Leu Gln Arg Leu Thr Gly Gly Cys Arg
	1365 1370 1375
30	Ala Arg Ala Ala Ala Val Val Val Gln Arg Gly Arg Arg Val Gly Trp
	1380 1385 1390
	Met Asp Gly Ser Trp Met Ser Leu Val Pro
	1395 1400